



UNIVERSIDAD AUTÓNOMA DE MADRID

FACULTAD DE CIENCIAS

DEPARTAMENTO DE BIOLOGÍA MOLECULAR

***Papel de la selección sexual post-cópula sobre la
evolución de genes reproductivos en mamíferos***

Alberto Vicens Sánchez

Madrid, 2014



UNIVERSIDAD AUTÓNOMA DE MADRID

FACULTAD DE CIENCIAS

DEPARTAMENTO DE BIOLOGÍA MOLECULAR

Papel de la selección sexual post-cópula sobre la evolución de genes reproductivos en mamíferos

Memoria presentada por Alberto Vicens Sánchez, licenciado en Biología.

Dirigida por Montserrat Gomendio, Profesora de Investigación del Museo Nacional de Ciencias Naturales (Consejo Superior de Investigaciones Científicas).

Tutelada por Cristina Grande Pardo, Profesora del departamento de Biología Molecular, Universidad Autónoma de Madrid.

Lugar de Realización de la Tesis:

Museo Nacional de Ciencias Naturales (CSIC)

Departamento de Ecología Evolutiva

Grupo de Ecología y Biología de la Reproducción



*A mis padres, por su apoyo y confianza durante
toda mi vida de estudiante.*

Agradecimientos

Y llegó el momento de escribir los agradecimientos. Uno de los momentos más felices y deseados por el doctorando sin duda. No sólo porque implica que ya he terminado de escribir la tesis (aunque confieso que todavía me quedan por pulir unos detalles a día de hoy), sino porque este es un rato emotivo, durante el cual echo la vista atrás, y recuerdo a todas aquellas personas (que en mi caso afortunadamente no son pocas) que he conocido y que me han aportado tantas y tan buenas cosas durante estos más de cuatro años.

En primer lugar, agradecer a Montse Gomendio, la persona que me abrió las puertas del mundo de la ciencia al darme la oportunidad de realizar el doctorado y que me guió durante mis primeros pasos por el camino de la investigación. Un camino apasionante pero que, sobre todo al principio, presenta grandes obstáculos. A Eduardo Roldán, porque sin sus enseñanzas, consejos, y sobre todo sus ánimos en los momentos más difíciles, probablemente no hubiera tenido las fuerzas suficientes para afrontar el gran reto de la tesis.

A Cristina Grande, mi tutora en la Universidad Autónoma, por su atención y por echarme una mano con los tediosos trámites burocráticos de la tesis.

Al Ministerio de Ciencia e Innovación (actualmente con el nombre de Ministerio de Economía y Competitividad), por haberme otorgado la beca predoctoral FPI para poder realizar la tesis.

A François Bonhomme y Annie Orth del Institut des Sciences de l'Évolution, CNRS-Université de Montpellier, por aportarnos los individuos de las especies de Mus que necesitábamos. También a Juan Luque-Larena (Juancho) por realizar las capturas de individuos salvajes y su aportación de ejemplares.

Agradecer incansablemente a Steve Dorus, mi supervisor en las dos estancias que realicé durante el doctorado. Sin duda una persona clave en mi carrera, cuya profesionalidad y su ejemplar manera de trabajar me hicieron crecer enormemente como investigador. Al mencionar a Steve, no puedo evitar acordarme de Alex Ball y Kirill Borziak, mis compañeros (y amigos) que tanto me ayudaron durante mis estancias en Inglaterra y Estados Unidos respectivamente.

Y ahora toca el turno de agradecer a toda la maravillosa gente que he conocido en el museo durante este tiempo.

En primer lugar, a todos mis compañeros de grupo. A los primeros que conocí y ya se fueron (Cris, Marta, Natalia, Mari Jose, Lucía, Juan, Noe y Valeria), que me enseñaron a hacer los primeros “pinitos” como científico y que me orientaron cuando estaba tan perdido al empezar. A este grupo también pertenecería Laura, pero mencionarla aparte porque ella me sigue ayudando día a día. A mis compañeros en el presente: a Lena, mi compañera de viaje durante la tesis, por su contribución a muchos de mis trabajos; a Maxi, por sus aportaciones, sus útiles consejos y sus interesantes charlas científicas, filosóficas y de baloncesto; a María, con quien he desahogado las desgracias típicas de un doctorando; a Javi, que siempre está allí y tira del grupo como nadie; a Pilar, por su compañía mientras nos fumamos “pitis” y por ayudarme con mis “rayadas”; a Dani, mi inseparable compañero durante el último año y medio y que en tantas cosas me ha apoyado; a Ester, mi “roldana” favorita, a la que dejo mi legado tanto dentro como fuera del museo; A Juan Antonio y Esperanza, por cuidar de nuestros “animalitos” y porque sin su labor los demás no podríamos trabajar.

Como no, agradecer a la gente del laboratorio de molecular (Piluchi, Isabel, Ricardo, Lourdes e Iván Acevedo), que tanto me han ayudado y tantas cosas me han enseñado durante mis largas horas allí dentro.

A mis colegas del Departamento de Ecología Evolutiva. Aquí tuve serias dudas sobre qué criterio aplicar para agrupar a la gente (“jaqueteros” o “tapperos”, “pestuzos” o “moñas”, “titulares” o “suplentes”...). Pero para evitar que alguien se pudiera ofender en un escrito tan emotivo, decidí basarme en la distribución por despachos. Así que ahí va. A los de la 1111: Carlos Ponce, Juan Navarro, Isaac, Octavio y Alex Ibañez (pintan los bastos). A los del “ciber”: Raúl, Chio, Ramón, Paco, Marga, Melinda, Silvia, Pau, Andrés, Valentina y Marcos (¿pues entonces qué coj...?). A los “merinos”: Elisa, Juan Rivero y Rigo. A los “giles”: Jaime, Blanca y Sergio. A los “morenos”: Alex Cantarero y Jimena. A los “avutarderos”: Aurora, Carlos Palacín e Iván Salcedo. A los “juanjos”: Eva y el doctor Barrientos. Al “pingüinero” Roger, que es un tío “apustuflant”. A Jose, por sus chistes que me alegraban los días. A Martí, porque no nos quedemos sin comer...

A la gente de Biodiversidad, que no son tan numerosos pero no por ello menos importantes. Miriam, Jorge, David, Patri, Juanes, Silvia, Guaci e Iker. Sois todos unos grandes, y aprovecho para recordar a los lectores que fuisteis vosotros los que instaurasteis las cañas de los Jueves, ¡Que no os embauquen!

Llegados a este punto, puede haber tres personas “mosqueadas” por pensar que me he olvidado de ellos. ¿Pero cómo olvidarme de mis hermanos del Campamento Krusty? Diego, Paloma y Carolina, los tres “liantes” del museo por excelencia. ¡Gracias por estar allí siempre!

Liquidada toda la gente del Museo, ahora me toca agradecer a todos los que me han apoyado desde fuera. A mis compañeros del Master de la Autónoma: Jero, Fer, Pablo, Andrea, Noe, Azu, Carlos, Elia, Curro, Ali, Álvaro, Arantxa, Raquel, María, Pachuca y Garavís. Como podéis ver, la tesis sobre “la reconstrucción del genoma del T-Rex” era una farsa, pero espero que esta no os disguste. A mis amigos de Usera (Pecha, Raúl, Velayos, Pela, Pablix y Julio) que se preocupaban constantemente por como llevaba la tesis. A mi gente de la Complu (Varas, Bea, Isa, Adri, Rocío, Olga, Álvaro, Esther y Elsa), y que vean que se pueden hacer tesis en cuatro años y medio y con buena nota.

Al margen de las personas que conozco, también quería dar las gracias a Bon Jovi, Kase-O, Arctic Monkeys, Extremoduro, The Doors, y otros muchos artistas por haber compuesto la banda sonora de esta tesis.

Aunque ya les dedico la tesis, no puedo olvidar de mencionar también aquí a mis padres, que desde que estaba en pre-escolar confiaron en mí y siempre estuvieron convencidos de que podría alcanzar grandes metas (aunque eso todavía está por ver...). También agradecer a mi hermano Rafa, simplemente por ser lo que es para mí.

Y por último, agradecer esta tesis por encima de todo, a un tal “Chechu” que, aunque no firma en ninguno de los trabajos, siento que es alguien que se ha currado esta tesis de principio a fin.

Nos vemos en el Asador... ¡y aupa Atleti!

“La belleza es el resultado de una selección sexual”

Charles Darwin

La investigación desarrollada en esta Tesis Doctoral ha sido financiada por la Dirección General de Investigación Científica y Técnica, Secretaría de Estado de Investigación, Desarrollo e Innovación, del Ministerio de Economía y Competitividad, a través de los proyectos CGL2008-04090 (IP: Montserrat Gomendio Kindelan) y CGL2011-26341 (IP: Eduardo Roldán Schuth), y de beca/contrato de Formación de Personal Investigador BES-2009-029239. Las estancias breves en la Universidad de Bath, Gran Bretaña y en Syracuse University, Estados Unidos, han sido también financiadas por el Ministerio de Economía y Competitividad (EEBB-I-2012-04676 y EEBB-I-2013-07479).



Algunos de los resultados obtenidos en esta tesis fueron incluidos en las siguientes publicaciones:

Vicens. A., Lüke. L., Roldan. E. R. S. 2014. Proteins involved in motility and sperm-egg interaction evolve more rapidly in mouse spermatozoa. PLoS One 9(3):e91302. doi: 10.1371/journal.pone.0091302.

Vicens, A., Tourmente, M., Roldan E. R. S. 2014. Structural evolution of Catsper1 in rodents is influenced by sperm competition, with effects on sperm swimming velocity. BMC Evolutionary Biology 14:106. doi: 10.1186/1471-2148-14-106.

Vicens, A., Roldan E. R. S. 2014. Coevolution of positively selected IZUMO1 and CD9 in rodents: Evidence of interaction between gamete fusion proteins? Biology of Reproduction 90(5):113. doi: 10.1095/biolreprod.113.116871.

Vicens, A., Gómez Montoto, L., Couso-Ferrer, F., Sutton, K.A., Roldan, E. R. S. 2014. Sexual selection and the adaptive evolution of PKDREJ protein in rodents. Molecular Human Reproduction (MHR-14-0095.R1: re-enviado-en revisión).

ÍNDICE

Resumen	1
Summary	2
Clave de abreviaturas	3
1. Introducción	5
1.1. Selección sexual: concepto y origen	6
1.2. Tipos de selección sexual.....	6
1.2.1. Selección sexual pre-cópula.....	6
1.2.2. Selección sexual post-cópula	6
1.3. La competición espermática como fuerza selectiva.....	7
1.3.1. Cantidad de espermatozoides	8
1.3.2. Viabilidad de los espermatozoides	8
1.3.3. Longevidad de los espermatozoides	8
1.3.4. Motilidad y velocidad de los espermatozoides	9
1.3.5. Tamaño y morfología del espermatozoide	9
1.3.6. Fluido seminal.....	10
1.4. Evolución de las proteínas reproductivas.....	10
1.4.1. Las proteínas reproductivas evolucionan rápidamente	10
1.4.2. Análisis de evolución adaptativa.....	11
1.4.3. Evolución adaptativa de las proteínas reproductivas	11
1.4.4. Evolución de las proteínas reproductivas en invertebrados	12
1.4.5. Evolución de las proteínas reproductivas en mamíferos	15
1.5. Proteómica de las proteínas reproductivas.....	21
1.6. Causas de la evolución adaptativa de las proteínas reproductivas.....	21
1.6.1. Resistencia a patógenos	21
1.6.2. Selección críptica de la hembra.....	22
1.6.3. Competición espermática.....	22

1.6.4.Conflicto sexual	23
1.7. Competición espermática y la evolución de proteínas reproductivas	23
1.8. Los roedores como modelo en estudios de competición espermática.....	25
2. Objetivos.....	27
3. Material y Métodos.....	29
3.1. Animales	30
3.1.1.Especies utilizadas	30
3.1.2.Mantenimiento de los animales.....	32
3.2. Obtención de muestras.....	32
3.2.1.Sacrificio y disección del animal	32
3.3. Análisis de biología molecular.....	33
3.3.1.Extracción de ácidos nucleicos	33
3.3.2.Amplificación de secuencias por PCR.....	33
3.3.3.Electroforesis en gel de agarosa.....	35
3.3.4.Purificación de fragmentos amplificados.....	36
3.3.5.Obtención y procesamiento de secuencias.....	36
3.4. Análisis filogenéticos.....	37
3.4.1.Selección del modelo de evolución.....	37
3.4.2.Construcción de árboles filogenéticos	37
3.5. Análisis evolutivos.....	38
3.5.1.Estimación de tasas evolutivas específicas para cada codon	38
3.5.2.Estimación de tasas evolutivas específicas para cada linaje	39
3.5.3.Estimación de tasas evolutivas específicas para cada linaje y entre codones.....	40
3.5.4.Análisis de coevolución	41
3.6. Medida de los niveles de competición espermática y de los parámetros espermáticos.....	41
3.6.1.Indice de competición espermática	41

3.6.2. Parámetros de velocidad espermática	42
3.7. Análisis de correlación corregidos por efectos filogenéticos.....	44
3.8. Análisis genómicos en cepas de ratón	44
3.8.1. Selección de proteínas.....	44
3.8.2. Obtención de secuencias de los genomas de ratón.....	44
3.8.3. Análisis estadísticos	45
3.9. Análisis de proteómica del espermatozoide.....	46
3.9.1. Preparación de muestras.....	46
3.9.2. Cuantificación de las muestras.....	46
3.9.3. Electroforesis diferencial en geles de 2 dimensiones (2D-DIGE)	46
3.9.4. Cromatografía líquida y espectrometría de masas cuantitativa.....	47
3.9.5. Cuantificación de proteínas con APEX.....	48
3.9.6. Clasificación funcional de proteínas del espermatozoide	49
4. Resultados.....	51
Bloque I: Las proteínas del espermatozoide de ratón implicadas en motilidad e interacción de gametos presentan una elevada tasa evolutiva.....	52
4.1. Distribución de proteínas	52
4.2. Tasas evolutivas de las proteínas del espermatozoide	53
4.3. Selección positiva en las proteínas del espermatozoide.....	56
Bloque II: La evolución estructural de Catsper1 en roedores está influida por la competición espermática, con efectos sobre la velocidad de natación del espermatozoide	59
4.4. Secuencias y análisis filogenéticos de Catsper1	59
4.5. Análisis de selección positiva en Catsper1	63
4.6. Asociación entre la evolución de Catsper1 y los niveles de competición espermática.....	65
4.7. Asociación entre la evolución de Catsper1 y la velocidad de natación espermática	67
4.8. Evolución de los residuos de histidina.....	69

Bloque III: Evolución de la proteína PKDREJ y el efecto de la competición espermática en roedores71

4.9. Secuencias y análisis filogenéticos de PKDREJ en roedores	71
4.10. Evidencia de selección positiva en PKDREJ de roedores	75
4.11. Asociación entre la evolución de PKDREJ y los niveles de competición espermática en roedores	77
4.12. Detección de selección positiva en linajes de roedores	78
4.13. Secuencias y análisis filogenéticos de PKDREJ en primates	83
4.14. Asociación entre la evolución de PKDREJ y los niveles de competición espermática en primates	84
4.15. Detección de selección positiva en linajes de primates.....	85

Bloque IV: Coevolución entre dos proteínas implicadas en la fusión de gametos: Izumo1 y Cd986

4.16. Secuencias y análisis filogenéticos.....	86
4.17. Análisis de selección positiva.....	91
4.18. Localización de los residuos bajo selección positiva en modelos estructurales.....	92
4.19. Coevolución entre Izumo1 y Cd9.....	93
4.20. Comparación de sustituciones de Izumo1 y Cd9 en los linajes.....	99
4.21. Asociación entre la evolución de Izumo1 y Cd9 y los niveles de competición espermática.....	99

Bloque V: Análisis de proteómica cuantitativa revelan diferencias en la abundancia de proteínas del espermatozoide entre especies de ratón con diferentes niveles de competición espermática.....100

4.22. Composición del proteoma del espermatozoide de ratón.....	100
4.23. Anotación funcional del proteoma del espermatozoide de ratón.....	102
4.24. Detección de proteínas con presencia exclusiva en las especies.....	105
4.25. Análisis cuantitativos de APEX.....	105
4.26. Análisis comparativos de los datos cuantitativos.....	107

4.27. Análisis de péptidos únicos.....	108
4.28. Cambios en abundancia de proteínas dirigidos por la competición espermática (análisis APEX).....	108
4.29. Análisis comparativos en geles 2D (2D-DIGE).....	111
4.30. Cambios en abundancia de proteínas dirigidos por la competición espermática (análisis 2D-DIGE).....	113
4.31. Comparación de los datos obtenidos entre los análisis de APEX y 2D-DIGE.....	114
5. Discusión.....	117
6. Conclusiones.....	131
Conclusiones en español.....	132
Conclusions in English	133
Bibliografía	135

RESUMEN

Se ha apreciado que las proteínas implicadas en reproducción evolucionan rápidamente en muchos grupos taxonómicos. En muchos casos, la rápida divergencia de estas proteínas es dirigida por la evolución adaptativa (selección positiva Darwiniana), lo que indica que los cambios fijados son beneficiosos para la reproducción. Para explicar la rápida evolución de las proteínas reproductivas, se han sugerido varias fuerzas evolutivas, incluyendo la selección post-cópula. La selección sexual post-cópula comprende procesos tales como la competición espermática, que es la competición entre los espermatozoides de dos o más machos por fecundar un óvulo. La competición espermática promueve la adaptación de diversos rasgos reproductivos con el fin de incrementar el éxito de fecundación. Se ha propuesto frecuentemente que los cambios adaptativos en rasgos reproductivos dirigidos por las fuerzas selectivas son el resultado de la rápida evolución de las proteínas reproductivas. Este trabajo se centra en el estudio de la evolución de las proteínas integrales del espermatozoide en roedores, con el objetivo de identificar proteínas que evolucionan bajo la influencia de la competición espermática.

Los diferentes estudios realizados revelaron que las proteínas con un papel en la motilidad espermática y en la interacción espermatozoide-óvulo muestran una mayor incidencia de evolución adaptativa, indicando que las fuerzas selectivas que actúan sobre el espermatozoide tienen un mayor impacto sobre estos dos procesos. Una proteína implicada en la regulación de la motilidad del espermatozoide, *Catsper1*, contiene un dominio intracelular que presenta una elevada tasa de inserciones y deleciones. Esta variación en longitud está promovida por selección positiva y la competición espermática favorece un acortamiento de esta región, lo cual se asocia con un incremento en la velocidad de natación del espermatozoide. Por otra parte, la proteína *PKDREJ*, que regula el transporte y la reacción acrosómica de los espermatozoides, presenta mayor divergencia en especies más promiscuas, mostrando que esta proteína puede ser importante para el éxito reproductivo en situaciones de competición espermática. Dos proteínas esenciales en la fusión de los gametos (*Izumo1* en el espermatozoide y *Cd9* en el óvulo) muestran selección positiva en dominios extracelulares implicados en la interacción con proteínas. Además, estos dos dominios están sujetos a un proceso de coevolución, sugiriendo que *Izumo1* y *Cd9* podrían tener una interacción directa durante la fecundación. Por último, los análisis de proteómica detectaron diferencias cuantitativas de proteínas del espermatozoide entre especies con diferentes niveles de competición espermática. Las especies de alta competición espermática presentan mayor abundancia de algunas proteínas con funciones importantes para el éxito reproductivo, revelando que la selección sexual puede promover cambios en la expresión de proteínas.

Estos resultados dan evidencia del papel importante que tienen las fuerzas de selección sexual, y en concreto la competición espermática, en la evolución de las proteínas reproductivas.

SUMMARY

It has been appreciated that proteins involved in reproduction are rapidly evolving in many taxonomical groups. In many cases, the rapid divergence of these proteins is driven by adaptive evolution (positive Darwinian selection), which indicates that the fixed changes are beneficial to reproduction. To explain the rapid evolution of reproductive proteins, several evolutionary forces have been identified, including post-copulatory sexual selection. Post-copulatory sexual selection comprises several processes such as sperm competition, which is the competition between the sperm of two or more males to fertilize the ova. Sperm competition promotes the adaption of diverse reproductive traits to increase the fertilization success. It has often been proposed that adaptive changes in reproductive traits driven by selective forces are the result of the rapid evolution of reproductive proteins. This work focuses on studying the evolution of integral sperm proteins in rodents with the purpose of identifying proteins evolving under sperm competition.

The different studies revealed that proteins with a role in sperm motility and sperm-egg interaction show a higher prevalence of adaptive evolution, indicating that the selective forces acting on spermatozoa have a higher impact on these two processes. A protein involved in the regulation of sperm motility, *Catsper1*, contains an intracellular domain showing an elevated rate of insertions and deletions. Length variation of this protein is promoted by positive selection, and sperm competition favors a shortening of this region, which is associated with increases in sperm swimming velocity. On the other hand, *PKDREJ* protein, which regulates sperm transport and acrosome reaction, shows higher divergence in more promiscuous species, revealing that this protein could be important for reproductive success under situations of sperm competition. Two essential proteins in gamete fusion (*Izumo1* in the spermatozoon and *Cd9* in the egg) show positive selection in the extracellular domains involved in protein interaction. In addition, these two domains are subjected to a coevolutionary process, suggesting that *Izumo1* and *Cd9* could have a direct interaction during fertilization. Finally, proteomic analyses detected quantitative differences in sperm proteins among species with different levels of sperm competition. High sperm competition species show higher abundance of some proteins with important functions for reproductive success, revealing that sexual selection is able to promote changes in protein expression.

Altogether, these results give evidence that sexual selective forces, and in particular sperm competition, have an important role in the evolution of reproductive proteins.

CLAVE DE ABREVIATURAS

2D-PAGE: Electroforesis en gel de poliacrilamida 2D (2D-PolyAcrilamide Gel Electrophoresis)

ADNc: ADN complementario

APEX: Expresión absoluta de proteínas (*Absolute Protein Expression*)

BEB: Análisis Bayseiano Empírico (*Bayesian Empirical Bayes*)

BLAST: Herramienta de búsqueda de alineamientos locales (*Blast Local Alignment Search Tool*)

DPBS: Suero salino de fosfato Dubecco (*Dubecco Phosphate-Buffered Saline*)

GO: Ontología de genes (*Gene Ontology*)

IEF: Enfoque isoelectrico (*IsoElectric Focusing*)

IUPAC: Unión internacional de química pura y aplicada (*International Union of Pure and Applied Chemistry*)

LC-MS/MS: Cromatografía líquida y Espectrometría de Masas (*Liquid Cromatography and Mass Spectrometry*)

LRT: Análisis de diferencia de Máxima Verosimilitud (*Likelihood Ratio Test*)

mT-H: Medio Tyrode modificado con Hepes

PAML: Análisis filogenéticos por Máxima Verosimilitud (*Phylogenetic Analyses by Maximum likelihood*)

PCA: Análisis de Componentes Principales (*Principal Component Analysis*)

PCR: Reacción en cadena de la polimerasa (*Polymerase Chain Reaction*)

PGLS: Análisis filogenético de mínimos cuadrados (*Phylogenetic Generalized Least Squares*)

PKDREJ: *Polycystic Kidney-Disease Receptor for Egg Jelly*

RTM: Masa relativa de testículo (*Relative Testis Mass*)

TBE: Tampón de electroforesis (compuesto por Tris, Borato y EDTA)

UTR: Región sin traducir (*UnTranslated Region*).

1. INTRODUCCIÓN

1.1. SELECCIÓN SEXUAL: CONCEPTO Y ORIGEN

Años después de la publicación de *El origen de las especies*, Charles Darwin escribió su segundo libro sobre la teoría de la evolución: *El origen del hombre, y la selección en relación al sexo* (1871). En esta obra, Darwin remarcó la importancia de la selección sexual en la evolución de las especies, y proponía la selección sexual como un proceso que opera sobre la capacidad de los machos para competir con otros machos con el fin de acceder a oportunidades reproductivas, lo que promueve el desarrollo de rasgos que confieran una ventaja en la competición reproductiva del individuo.

1.2. TIPOS DE SELECCIÓN SEXUAL

1.2.1. Selección sexual pre-cópula

Durante muchas décadas la selección sexual se estudió en especies monógamas o poligínicas, en las que los espermatozoides de un solo macho son los únicos disponibles para fecundar el óvulo de una hembra determinada. De este modo, el macho que copula tiene la garantía de transmitir su linaje genético a toda la descendencia generada por esa hembra. En estos sistemas de apareamiento, los machos se ven obligados a competir entre ellos por reproducirse con las hembras limitadas de su entorno, por lo que la selección sexual opera exclusivamente antes de la cópula en un proceso conocido como selección sexual pre-cópula (Fig. 1) (Andersson 1994). La selección sexual pre-cópula puede actuar a través de dos mecanismos: selección intrasexual y selección intersexual (Fig. 1). La selección intrasexual implica, en la mayoría de los casos, una rivalidad entre los machos por acceder a las hembras, y por ello este proceso se le suele denominar competición entre machos. La selección intersexual es un proceso por el que, normalmente las hembras, eligen a su pareja en función de los beneficios que pueden aportar los machos, y también se conoce como elección de la hembra.

La competición entre machos y la elección de la hembra son mecanismos de selección sexual igualmente importantes, y ambos pueden estar presentes de manera simultánea en las especies (Darwin 1871)

1.2.2. Selección sexual post-cópula

En un gran número de especies se ha observado que las hembras pueden aparearse con múltiples machos dentro del mismo ciclo reproductivo (es decir, son poliándricas). Como consecuencia de esto, el eyaculado de varios machos va a coincidir en el tracto reproductivo de la hembra y los gametos de diferentes individuos tendrán que competir para fecundar el óvulo (Birkhead y Moller 1998). Por tanto, la selección sexual va a proseguir incluso después de la cópula y la inseminación. Este fenómeno se conoce como selección sexual post-cópula y tiene lugar a través

de dos posibles mecanismos: la competición espermática (selección intrasexual) y la elección críptica de la hembra (selección intersexual) (Fig. 1). La competición espermática es una fuerza que promueve rasgos que incrementan el éxito de fecundación de un eyaculado cuando compite con el de otros individuos (Parker 1970). La selección críptica de la hembra incluye todos los mecanismos por los que la hembra es capaz de discriminar entre los espermatozoides de diferentes individuos y seleccionar el de un macho determinado para la fecundación (Thornhill y Alcock 1983; Eberhard 1996).

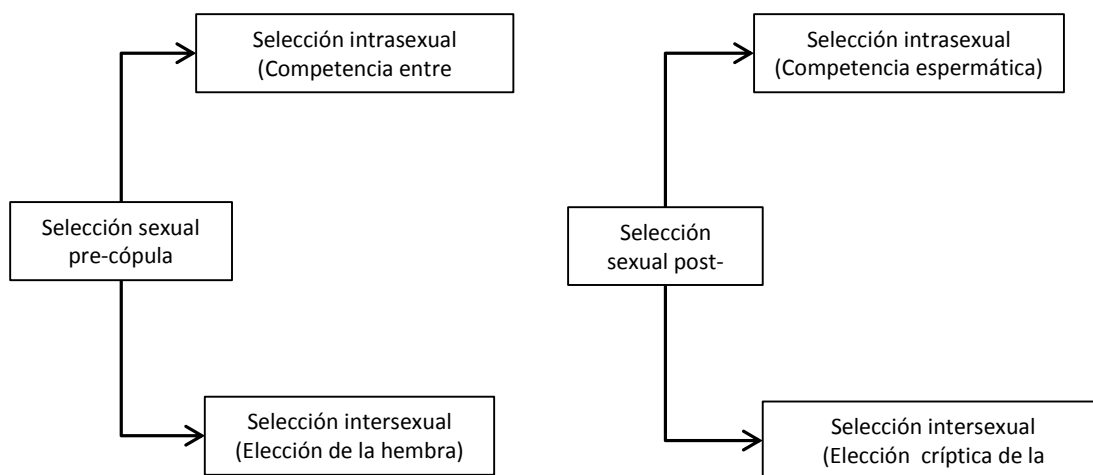


Figura 1. Diagrama representando los principales mecanismos de selección sexual.

1.3. LA COMPETICIÓN ESPERMÁTICA COMO FUERZA SELECTIVA

La competición espermática fue discutida inicialmente por G. A. Parker, quien la definió como la competición entre el esperma de dos o más individuos por la fecundación del óvulo (Parker 1970). La competición espermática es desde entonces considerada una importante y generalizada fuerza evolutiva, y se han llevado a cabo numerosos estudios en diversos taxones con el objetivo de desvelar el papel que tiene la competición espermática sobre la evolución del comportamiento reproductivo, la fisiología, la morfología y anatomía de los órganos reproductores, así como del fenotipo de los gametos (Smith 1984; Birkhead y Moller 1992; Simmons 2001; Shackleford y Pound 2006; Pizarri y Parker 2009). Los rasgos del eyaculado más importantes en la competición espermática se describen a continuación.

1.3.1. Cantidad de espermatozoides

En situaciones de competición espermática entre machos rivales, el número de espermatozoides transferidos en el eyaculado representa una ventaja importante, ya que al haber una mayor proporción de espermatozoides de un determinado macho en el tracto de la hembra, la probabilidad de fecundación de ese macho se verá incrementada (Parker 1970, 1984; Birkhead y Moller 1998; Parker y Pizarri 2010).

La selección de un mayor número de espermatozoides conlleva un aumento del tamaño del testículo, ya que las células espermáticas son originadas en los túbulos seminíferos de los testículos. Esta respuesta se ha observado en taxones tan diversos como insectos (Gage 1994), peces (Stockley *et al.* 1997), anfibios (Jennions y Passmore 1993), aves (Briskie y Robert 1992) y mamíferos (Gomendio *et al.* 1998), donde las especies expuestas a competición espermática presentan un mayor tamaño de testículos en relación al tamaño corporal. Por lo tanto, el peso relativo de testículos se ha utilizado generalmente como un índice representativo de los niveles de competición espermática entre las especies.

1.3.2. Viabilidad de los espermatozoides

Dado que una alta proporción de espermatozoides muere a su paso por el tracto genital de la hembra, por las barreras anatómicas y fisiológicas que este impone, la viabilidad del espermatozoide es importante para el éxito reproductivo (Gomendio *et al.* 1998; Scott 2000).

Se ha observado que las especies y poblaciones de insectos poliándricas tienen una mayor proporción de espermatozoides vivos que aquellas monoándricas (Hunter y Birkhead 2002; Garcia-Gonzalez y Simmons 2005). No obstante, un análisis reciente en roedores no detectó una correlación significativa entre niveles de competición espermática y viabilidad espermática cuando este rasgo fue comparado de manera aislada, aunque sí mostró una asociación significativa cuando fue analizado junto con otros rasgos determinantes de la calidad del eyaculado (Gómez Montoto *et al.* 2011a).

1.3.3. Longevidad de los espermatozoides

La longevidad puede estar determinada bien por la tasa metabólica del espermatozoide, o bien por la tasa de senescencia, que implica un mayor riesgo de mortalidad y por tanto una reducción en la capacidad de fecundación. La teoría predice que los espermatozoides más longevos estarán representados por más tiempo en el tracto femenino, por lo que las probabilidades de fecundación de un macho determinado aumentarán (Pizzari y Parker 2008).

Se han encontrado asociaciones entre longevidad y competición espermática en especies con fecundación interna, tanto a nivel interespecífico (Hunter y Birkhead 2002) como intraespecífico (Garcia-Gonzalez y Simmons 2005). Sin embargo, en especies con fecundación externa como los

peces, en los que la motilidad del espermatozoide es un componente crucial en la eficiencia de fecundación, la longevidad puede verse disminuida en favor de una motilidad más intensa (Gage *et al.* 2004). Los componentes reproductivos de la hembra pueden tener efectos sobre la longevidad de los espermatozoides (Turner y Montgomerie 2002).

1.3.4. Motilidad y velocidad de los espermatozoides

La motilidad espermática va a ser importante para el éxito en la fecundación por dos motivos. Por un lado, un eyaculado con una mayor proporción de espermatozoides móviles será más competitivo ya que es muy improbable que los espermatozoides sin motilidad lleguen a fecundar (Krzanowska 1974). Se ha encontrado evidencia de una asociación entre el porcentaje de espermatozoides móviles en el eyaculado y la intensidad de competición espermática en roedores (Gómez Montoto *et al.* 2011a). Por otro, es lógico asumir que los espermatozoides más rápidos alcanzarán el óvulo más pronto, por lo que la velocidad del espermatozoide va a ser un rasgo esencial en la competición espermática. Se ha demostrado que la velocidad espermática resulta un factor determinante para el éxito de fecundación tanto en condiciones competitivas (Birkhead *et al.* 1999; Gage *et al.* 2004; Gomendio y Roldan 2008; Gasparini *et al.* 2010) como no competitivas (Malo *et al.* 2005).

1.3.5. Tamaño y morfología del espermatozoide

Un buen número de trabajos han demostrado que los niveles de competición espermática se asocian con incrementos en el tamaño (normalmente cuantificado como longitud) del espermatozoide en grupos taxonómicos muy diversos (Gage 1994; Byrne *et al.* 2003; Gomendio y Roldan 2008; Fitzpatrick *et al.* 2009; Lüpold *et al.* 2009a; Tourmente *et al.* 2009, 2011b). Por otra parte, en especies donde los espermatozoides se almacenan en órganos especializados del tracto femenino, se ha observado que el tamaño del espermatozoide está más determinado por el tamaño del órgano de almacenaje (Miller y Pitnick 2002; García-González y Simmons 2007; Immler *et al.* 2007). Por tanto, la medida en que el tamaño del espermatozoide afecta a la eficiencia de fecundación parece variar entre diferentes grupos taxonómicos.

La motilidad del espermatozoide pueda verse afectada en diferente medida dependiendo del componente del espermatozoide que es modificado. El tamaño y la forma de la cabeza del espermatozoide pueden ser rasgos importantes porque determinan el grado de resistencia al avance, por lo que un alargamiento de la cabeza reduciría tal resistencia y resultaría en un incremento de la velocidad de natación (Malo *et al.* 2006; Tourmente *et al.* 2011b). Un aumento en el tamaño de la pieza media puede suponer un mayor suministro de energía para el movimiento del espermatozoide, dado que es el componente que contiene las mitocondrias (Anderson y Dixon 2002; Anderson *et al.* 2005; Tourmente *et al.* 2009). Un aumento en la longitud del resto del flagelo (pieza principal) también puede ser importante porque puede incrementar la fuerza de propulsión del espermatozoide,

además de generar una mayor cantidad de energía mediante la glicólisis (Malo *et al.* 2006; Gomendio y Roldan 2008; Tourmente *et al.* 2011a,b). Se ha sugerido que la relación de tamaño entre los diferentes componentes del espermatozoide también puede influir en la velocidad del espermatozoide (Malo *et al.* 2006; Humphries *et al.* 2008; Tourmente *et al.* 2011a,b).

1.3.6. Fluido seminal

En la mayoría de las especies, muchas sustancias, aparte de los espermatozoides, son transferidas en el eyaculado durante la inseminación. Muchas de las acciones de estas sustancias van a afectar directamente al espermatozoide, pudiendo modificar su calidad y transporte (Eberhard 1996). Además, los productos del fluido seminal son capaces de influir en las respuestas fisiológicas y en el comportamiento de las hembras (Simmons 2001). Por lo tanto, la acción de los diversos componentes del fluido seminal puede ser importante para el éxito reproductivo de los machos.

El papel de péptidos del fluido seminal en la competición espermática ha sido ampliamente estudiado en insectos (Chapman 2001; Simmons 2001; Wolfner 2002; Hodgson y Hosken 2006; Poiani 2006; Cameron *et al.* 2007), aunque también se han identificado algunas proteínas seminales importantes para la competición espermática en mamíferos (Jensen-Seaman y Li 2003; Ramm *et al.* 2009).

1.4. EVOLUCIÓN DE LAS PROTEÍNAS REPRODUCTIVAS

1.4.1. Las proteínas reproductivas evolucionan rápidamente

Como se ha comentado muchos rasgos del comportamiento, morfológicos, fisiológicos y celulares relacionados con la reproducción están expuestos a la selección sexual. Dado que los genes no reproductivos (aquellos que afectan al desarrollo, viabilidad y supervivencia del individuo) experimentan cambios adaptativos que son fijados por la selección natural, se propuso que los genes con funciones en la reproducción y el sexo podrían evolucionar con gran rapidez, dirigidos por fuerzas de selección sexual (Singh y Kulathinal 2000).

Se definen como genes reproductivos todos aquellos genes que codifican para proteínas que intervienen en la reproducción sexual y que participan en la producción, maduración, almacenaje, función, transporte, reconocimiento y fecundación de los gametos. Por tanto, a estas proteínas se las denomina proteínas reproductivas.

Durante las dos últimas décadas ha habido un número creciente de estudios demostrando que las proteínas reproductivas divergen más rápidamente que las proteínas expresadas en tejidos no reproductivos (Singh y Kulathinal 2000; Swanson y Vacquier 2002; Clark *et al.* 2006; Turner y Hoekstra 2008). Un descubrimiento importante fue el de Civetta y Singh (1995), quienes utilizando la técnica de electroforesis en 2 dimensiones revelaron que las proteínas expresadas en tejidos

reproductivos de *Drosophila* son significativamente más divergentes que las proteínas de tejidos no reproductivos. Posteriormente, en un estudio en el que se compararon hasta 1880 genes ortólogos entre ser humano y roedores, se observó que entre el 10% de los genes con mayor divergencia había una alta presencia de genes reproductivos (Makalowski y Boguski 1998).

Teniendo en cuenta que las proteínas reproductivas regulan procesos esenciales para el *fitness* y la fertilidad de los individuos, su elevada divergencia resultó sorprendente y en seguida se adoptó la idea de que los genes reproductivos cambian rápidamente como consecuencia de una evolución adaptativa.

1.4.2. Análisis de evolución adaptativa

La estimación de la tasa evolutiva de un gen usando porcentajes de divergencia no aporta información sobre las causas potenciales de su rápida evolución. Por ejemplo, un gen podría acumular mutaciones constantes por una falta de limitación funcional. (Vanin 1985). Alternativamente, la rápida evolución de un gen puede ser consecuencia de evolución adaptativa, que tiene lugar cuando una serie de cambios en la proteína son fijados por la selección natural debido a que aportan alguna ventaja para el individuo (Swanson y Vacquier 2002).

Para poder conocer el perfil evolutivo de una proteína, se ha utilizado un método basado en comparar las secuencias de ADN que codifican para la proteína entre dos o más especies. Cada mutación identificada en el alineamiento se clasifica entonces como no sinónima (si modifica el aminoácido para el que codifica) o como sinónima (si no altera la secuencia de la proteína) y se definen las tasas de cambios no sinónimos (d_N) y sinónimos (d_S) (Yang y Bielawski 2000). d_N y d_S pueden ser directamente comparados calculando el cociente $\omega = d_N/d_S$, que se utilizará para inferir el perfil evolutivo de una proteína (Yang y Bielawski 2000). Si ω tiene un valor menor que 1 implica que los cambios deletéreos de una proteína están siendo eliminados por selección purificante. Un ω aproximado a 1 indica evolución neutral y que la proteína acumula mutaciones aleatorias debido a una relajación de la selección. Por otra parte, si el cociente ω es superior a 1, esto significa que existen cambios de aminoácidos que se fijan en una mayor frecuencia que los cambios sinónimos porque probablemente aportan una ventaja adaptativa. Por tanto, un valor de ω significativamente mayor que 1 da evidencia de que la selección positiva (o selección Darwiniana) está actuando sobre una proteína.

1.4.3. Evolución adaptativa de las proteínas reproductivas

Numerosos análisis durante los últimos años han mostrado que la rápida divergencia de los genes reproductivos es promovida por evolución adaptativa (o selección positiva), lo que refleja que los cambios producidos en las secuencias son ventajosos para la reproducción (Swanson y Vacquier 2002; Clark *et al.* 2006; Turner y Hoekstra 2008a). Esta observación por tanto apoya la existencia de una selección sexual promoviendo variación en el genoma reproductivo de los dos sexos.

Se ha detectado evidencia de proteínas reproductivas experimentando evolución adaptativa en grupos taxonómicos de eucariotas muy diversos y que participan en diferentes etapas de la reproducción (Swanson y Vacquier 2002; Clark *et al.* 2006). Las siguientes secciones se centran en la divergencia de proteínas implicadas en reproducción animal. No obstante, en otros grupos como protozoos (Luporini *et al.* 1995), algas (Ferris *et al.* 1997; Armbrust y Galindo 2001), hongos (Casseltan y Olesnick 1998) y plantas (Richman y Kohn 2000; Kachroo *et al.* 2001) se han identificado proteínas reproductivas que también evolucionan con gran rapidez.

1.4.4. Evolución de las proteínas reproductivas en invertebrados

Los sistemas reproductivos que se han estudiado con mayor detalle en las últimas décadas han sido los de los invertebrados marinos como los abulones o los erizos de mar, y se han identificado pares de proteínas gaméticas interaccionando exclusivamente en estos dos grupos taxonómicos (Swanson y Vacquier 1997; Zigler 2008).

Los abulones son especies de moluscos del género *Halotis*. Esta especie presenta sexos separados, y ambos liberan sus gametos al agua del mar, por lo que su fecundación es externa (Lindberg 1992). Una vez que un espermatozoide de abulón encuentra un óvulo, nada a través de la primera barrera de protección, que es una capa gelatinosa. Una vez superada la capa gelatinosa, el espermatozoide alcanza la cubierta vitelina (CV), que es una envoltura más gruesa formada por una matriz de glicoproteínas entrelazadas (Kresge *et al.* 2001; Kamei y Glabe 2003). El contacto entre el espermatozoide y la CV induce la reacción acrosómica, un proceso que implica la fusión de las membranas externa e interna del acrosoma, un gránulo secretorio localizado en la parte apical de la cabeza del espermatozoide (Lewis *et al.* 1982). La reacción acrosómica promueve la liberación de una serie de proteínas del acrosoma que van a desasociar las fibras de la CV, formando un poro por el que va a pasar el espermatozoide. Una vez el espermatozoide atraviesa la CV, se une y se fusiona con la membrana del óvulo para ser transferido al citoplasma (Kresge *et al.* 2001).

Es frecuente que diferentes especies de abulones convivan en el mismo hábitat y que presenten etapas reproductivas solapantes. No obstante, se ha observado que estas especies presentan especificidad en su fecundación (Leighton y Lewis 1982), por lo que se planteó que el resultado de la fecundación es determinado por el reconocimiento específico de las moléculas expresadas en la superficie de los gametos.

La fecundación del abulón es controlada en gran medida por las interacciones entre las proteínas lisina en el espermatozoide y VERL en el óvulo (Swanson y Vacquier 1997). VERL es una glicoproteína larga con 22 repeticiones en tándem y constituye el principal componente de la CV (Swanson y Vacquier 1997). Por su parte, la lisina es una proteína que, tras la reacción acrosómica, se expone a la CV, donde disuelve las fibras de VERL mediante reacciones no enzimáticas, formando el poro para que el espermatozoide penetre a través de la CV (Lewis *et al.* 1982; Swanson y Vacquier

1997). La interacción entre la lisina y la CV es altamente específica, contribuyendo al aislamiento reproductivo entre especies (Swanson y Vacquier 1997). La lisina ha sido secuenciada en una gran cantidad de especies de abulones y se ha observado que esta proteína presenta una alta divergencia promovida por selección positiva (Lee *et al.* 1995). La lisina presenta dos regiones hipervariables que determinan la especificidad de la interacción con la CV (Lyon y Vacquier 1999), y dentro de estos dominios se ha observado una concentración de residuos bajo selección positiva (Yang *et al.* 2000b).

Otra proteína que se libera del espermatozoide del abulón durante la reacción acrosómica es la Sp18, que se cree que está implicada en la fusión del espermatozoide con el óvulo. La Sp18 es la proteína que evoluciona más rápidamente de las descritas hasta ahora en metazoos (Swanson y Vacquier 1995a).

Se ha encontrado evidencia de selección positiva para VERL en las dos primeras repeticiones, aunque no en el resto (Galindo *et al.* 2003). Estas 2 repeticiones podrían ser las responsables de interaccionar específicamente con la lisina ya que son las únicas que evolucionan de manera independiente (Galindo *et al.* 2003), pues las repeticiones de la 3 a la 22 son muy similares y están sometidas a un proceso de homogeneización por conversión génica o sobre cruzamiento desigual (Armbrust y Galindo 2001).

La evolución adaptativa de las regiones implicadas en la interacción lisina-VERL sugiere que estas proteínas de fecundación podrían estar sujetas a un proceso de coevolución con el fin de mantener el reconocimiento espermato-óvulo específico de cada especie (Fig. 2). De hecho, estudios recientes han mostrado evidencias que apoyan la hipótesis de que lisina y VERL coevolucionan en especies de abulones (Clark *et al.* 2009) y otras especies de moluscos marinos del género *Tegula* (Hellberg *et al.* 2012).

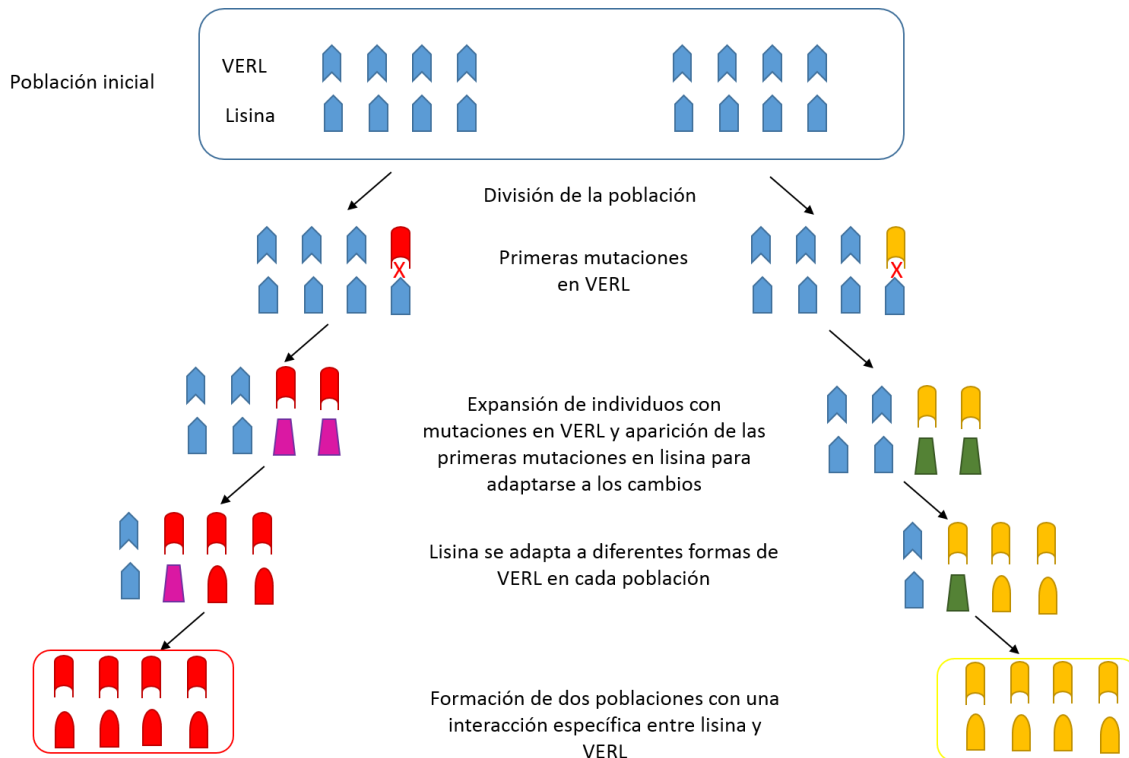


Figura 2. Modelo de coevolución de lisina y VERL. En diferentes poblaciones VERL puede sufrir diferentes tipos de cambio. En cada población, la lisina cambiará para adaptarse a las nuevas formas de VERL y poder mantener la interacción. Tras varias generaciones, las poblaciones fijan diferentes formas de lisina y VERL, por lo que la interacción gamética entre individuos de diferentes poblaciones será menos eficiente, dando lugar a un aislamiento reproductivo.

El género de erizo de mar *Echinometra* presenta un número elevado de especies simpátricas, y entre estas se ha observado una alta especificidad en la fecundación (Metz *et al.* 1994). La especificidad del reconocimiento esperma-óvulo en estas especies está determinada por la “bindina”, que es la única proteína contenida en el acrosoma del erizo de mar y está implicada tanto en la unión a la cubierta vitelina como en la fusión entre el espermatozoide con el óvulo (Vacquier y Moy 1977). El receptor de bindina en la superficie del óvulo es la proteína EBR1 (Kamei y Glabe 2003). La unión entre bindina y EBR1, como ocurre con el sistema lisina-VERL en abulones, es específica para cada especie (Glabe y Lennarz 1979; Kamei y Glabe 2003). La bindina presenta gran divergencia cuando es comparada a nivel interespecífico, y en algunos géneros se ha detectado evidencia de selección positiva (Metz y Palumbi 1996; Zigler *et al.* 2003). Esta divergencia adaptativa de bindina se ha asociado con incompatibilidad gamética; las especies de erizos de mar con alelos similares de bindina son compatibles para fecundación, mientras que las especies con alelos más divergentes son incompatibles (McCartney y Lessios 2004; Zigler *et al.* 2005). Estudios intraespecíficos han revelado que la compatibilidad gamética entre individuos de la misma especie también puede ser predicha por el genotipo de bindina, de tal modo que hembras y machos con los mismos alelos de bindina tienen más probabilidad de fecundar entre sí que con individuos de distinto genotipo (Palumbi 1999;

Levitan y Ferrell 2006). El receptor EBR1, por su parte, también muestra elevada divergencia entre especies (Kamei y Glabe 2003). Estos estudios han sugerido que los genes codificando bindina y EBR1 deben estar ligados en el mismo cromosoma, aunque todavía no se ha demostrado.

Los estudios en abulones y erizos de mar, por tanto, han servido para demostrar que la rápida evolución de proteínas implicadas en fecundación va a tener un papel clave en el mantenimiento del reconocimiento gamético así como en el aislamiento reproductivo de las poblaciones.

Algunas especies de invertebrados no marinos también presentan rápida evolución de proteínas reproductivas. El ejemplo más caracterizado son los componentes secretados por las glándulas accesorias de machos de *Drosophila*, que presentan funciones muy diversas tales como inducir la ovulación de la hembra, reducir su receptividad a otros machos, formar el tapón copulatorio, mediar el almacenaje del esperma y proteger a los gametos de ataques microbianos (Wolfner 1997, 2002). Se ha observado que las proteínas de las glándulas accesorias muestran altos niveles de divergencia y que una elevada proporción presenta evidencia de selección positiva (Swanson *et al.* 2001a). La evolución de algunas de estas proteínas se ha estudiado en detalle. Un ejemplo es el gen que codifica para la proteína Acp26Aa, que induce la ovulación. El gen *Acp26Aa* es uno de los genes más divergentes del genoma de *Drosophila* y evoluciona bajo selección positiva (Tsaur y Wu 1997).

1.4.5. Evolución de las proteínas reproductivas en mamíferos

Se ha encontrado evidencia de rápida divergencia y selección positiva en proteínas implicadas en muchas de las etapas de la reproducción en mamíferos, que es el grupo de vertebrados sobre el que se han desarrollado la inmensa mayoría de los estudios sobre evolución de proteínas reproductivas (Swanson y Vacquier 2002; Clark *et al.* 2005; Turner y Hoekstra 2008a).

Además de en *Drosophila*, la evolución adaptativa de proteínas del fluido seminal ha sido estudiada también en mamíferos. Un análisis de 161 proteínas seminales en primates reveló que más del 10% evolucionan rápidamente, y siete de ellas además mostraron señales significativas de selección positiva (Clark y Swanson 2005). Estudios en roedores han revelado que el proteoma de la vesícula seminal presenta una composición muy diversa entre especies y que las proteínas expresados en esta glándula accesoria presentan una alta divergencia molecular y estructural (Dean *et al.* 2009; Ramm *et al.* 2009). Gran parte de los estudios evolutivos se han centrado en las semenogelinas, una familia de proteínas que forman un coágulo seminífero tras la eyaculación. En particular, se ha detectado evidencia de selección positiva en los genes *SEMG1* y *SEMG2* de primates, cuyo productos son los componentes más abundantes del coágulo seminífero (Kingan *et al.* 2003; Dorus *et al.* 2004) (Fig. 3). El ortólogo de *SEMG2* en roedores, *Svs2*, también muestra evidencia de evolución adaptativa (Ramm *et al.* 2008). Otro producto del fluido seminal que exhibe una señal de selección positiva es el codificado por el gen *Sva* (Ramm *et al.* 2008), que se une directamente al esperma,

actuando como un modulador de la capacitación (Huang *et al.* 2007), un proceso que ocurre en el oviducto y que confiere al espermatozoide la capacidad para fecundar el óvulo (Visconti y Moore 1998; Jaiswal y Eisenbach 2002)

Los análisis evolutivos del transcriptoma del epidídimo, que es el tejido donde maduran los espermatozoides tras diferenciarse en los túbulos seminíferos (Quill y Garbers 2002) han revelado que los genes que expresan productos secretados por el epidídimo muestran una señal de evolución adaptativa, mientras que los genes expresados en tejido pero no secretados se encuentran más conservados (Dean *et al.* 2008).

Durante las últimas etapas de la espermatogénesis, las histonas que se encuentran unidas inicialmente a la cromatina de las espermátidas son reemplazadas por las protaminas (Oliva y Dixon 1991). Las protaminas son una familia de proteínas ricas en arginina que se unen con mayor afinidad al ADN, consiguiendo una máxima compactación del núcleo del espermatozoide y pudiendo influir en la forma de la cabeza (Balhorn 2007). En algunos grupos de mamíferos se expresan dos tipos de protaminas (PRM1 y PRM2), y en estos se ha observado que tanto PRM1 y PRM2 muestran una rápida tasa de cambio (Retief *et al.* 1993; Queralt *et al.* 1995) (Fig. 3). Posteriormente, en primates se demostró que la rápida divergencia de PRM1 y PRM2 es resultado de episodios de selección positiva (Rooney y Zhang 1999; Wyckoff *et al.* 2000). En roedores, sólo se detectó una débil señal de selección positiva en PRM2 al comparar un grupo de muridos próximos filogenéticamente (Martín-Coello *et al.* 2009). No obstante, cuando las protaminas fueron comparadas entre especies de cricétidos más distantes, se observó que PRM1 muestra evidencia de selección positiva, mientras que PRM2 evoluciona bajo relajación funcional (Lüke *et al.* 2011). Por otra parte, se ha presentado evidencia de selección positiva para DAZLA en primates (Bielawski y Yang 2001), un gen cuyo producto es esencial para la gametogénesis (Ruggiu *et al.* 1997)

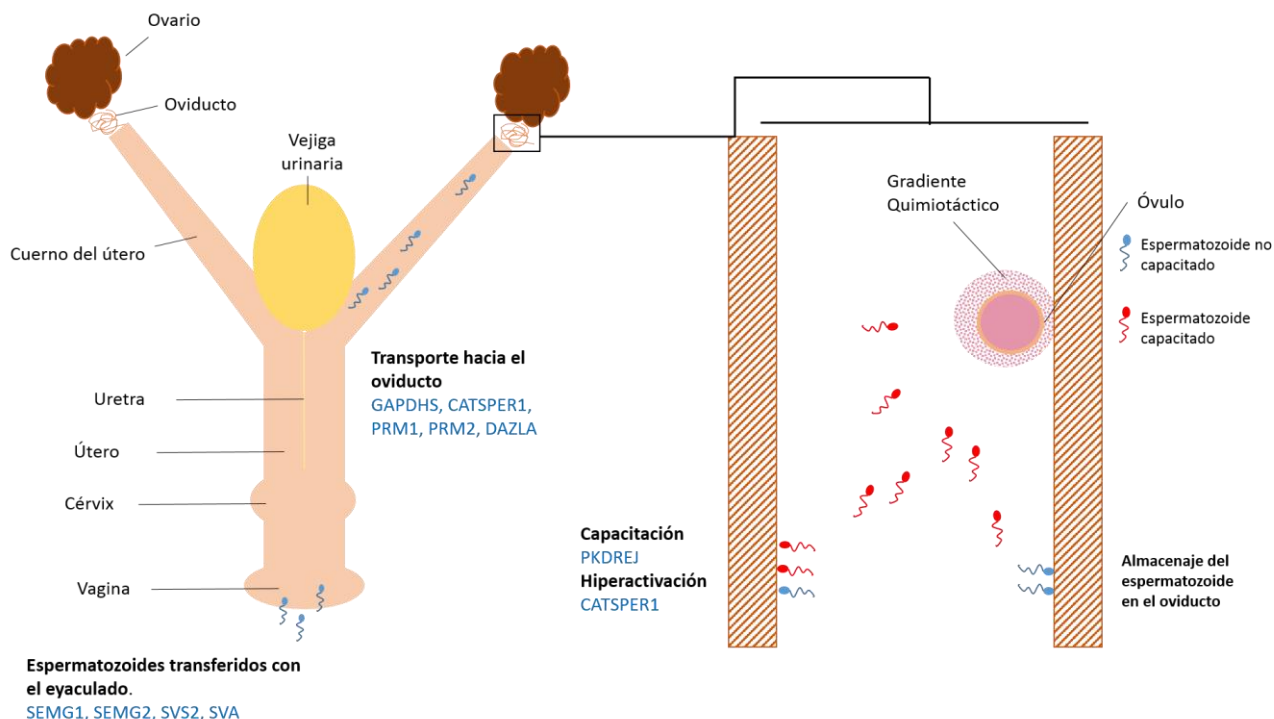


Figura 3. Representación del tracto reproductivo de ratón y de los procesos que experimentan los espermatozoides desde que son inseminados hasta que llegan al óvulo. Las proteínas con evidencia de selección positiva se indican en cada proceso reproductivo.

La motilidad espermática es un factor crucial en el éxito reproductivo, ya que el desplazamiento del espermatozoide le va a permitir avanzar a través del tracto reproductivo de la hembra. Estudios de proteómica en diversas especies de mamíferos han aportado un amplio catálogo de las proteínas expresadas en las diferentes estructuras del flagelo (Cao *et al.* 2006; Baker *et al.* 2007; 2008a,b; Skerget *et al.* 2013). No obstante, poco se conoce de la función de las proteínas del flagelo (Escalier 2006; Turner 2006; Escalier y Touré 2012). Dado que gran parte de las proteínas que regulan la motilidad van a tener funciones estructurales, metabólicas o de transporte, se espera que se mantengan altamente conservadas. No obstante, se ha observado que algunas proteínas implicadas en motilidad muestran una rápida divergencia. El gen *Catsper1* codifica para un canal que regula la entrada de calcio (Ca^{2+}) en el flagelo y es requerido para la motilidad y la hiperactivación del espermatozoide (Ren *et al.* 2001; Carlson *et al.* 2003; Qi *et al.* 2006; Ho *et al.* 2009) (Fig. 3). La hiperactivación es una forma de motilidad adquirida durante la capacitación, caracterizada por un batido más vigoroso del flagelo y con una mayor curvatura (Yanagimachi 1968; Suarez y Ho 2003a). La adquisición de la hiperactivación permite al espermatozoide desasociarse de la pared del oviducto, potenciar la motilidad para progresar a través del oviducto y penetrar las diferentes capas que rodean al óvulo (Suarez 2008). La proteína *Catsper1* presenta un dominio amino-terminal que experimenta un gran número de inserciones y deleciones en roedores y primates, dando lugar a una enorme variación en el tamaño de esta región (Podlaha y Zhang 2003; Podlaha *et al.* 2005). Además de *Catsper1*, otras proteínas de la familia *Catsper* también experimentan una

rápida evolución en mamíferos (Cai y Clapham 2008) Por otro lado, se ha observado evidencia de selección positiva en *Gapdhs* (Torgerson *et al.* 2002) (Fig. 3), un gen que codifica para una isoenzima implicada en la glicólisis del espermatozoide, por lo que las fuerzas selectivas también podrían actuar sobre proteínas que regulan la producción de energía para la motilidad espermática (Welch *et al.* 2000).

Una vez que el espermatozoide alcanza el óvulo, la estructura y la secuencia de las proteínas implicadas en el reconocimiento gamético puede determinar el resultado de la fecundación entre individuos de la misma especie. Aunque se ha identificado un gran número de proteínas tanto en la superficie del espermatozoide como en la del óvulo, a día de hoy no se han identificado parejas de proteínas que interaccionan entre sí en mamíferos. Existen multitud de proteínas expresadas en la superficie de los gametos que muestran tasas de evolución muy altas (Swanson *et al.* 2001b, 2003; Glassey y Civetta 2004; Herlyn y Zischler 2008). Un reciente estudio de proteómica reveló que una alta proporción de proteínas expresadas en la membrana del espermatozoide de mamíferos presentan señales de selección positiva (Dorus *et al.* 2010).

El espermatozoide en su aproximación al óvulo tiene que atravesar una matriz formada por las células del cúmulo, y durante este tránsito iniciará la reacción acrosómica (Hong *et al.* 2009; Jin *et al.* 2011). El espermatozoide entonces se unirá a la zona pelúcida, una envuelta del óvulo formada por glicoproteínas, para finalizar el proceso de reacción acrosómica (Wassarman *et al.* 2001). En la zona pelúcida de mamíferos se han identificado hasta cuatro glicoproteínas (ZP1, ZP2, ZP3 y ZP4), y las especies pueden presentar tres o cuatro de estas glicoproteínas en función del grupo (Goudet *et al.* 2008). ZP2 y ZP3 son las glicoproteínas más abundantes en la zona pelúcida y se ensamblan constituyendo una matriz (Wassarman *et al.* 2004) (Fig. 4). Mientras que a ZP2 se le ha atribuido una función como receptor secundario al que se une el espermatozoide tras la exocitosis del acrosoma (Wassarman *et al.* 2004), ZP3 es el principal receptor del espermatozoide y se la ha considerado como una proteína inductora de la reacción acrosómica (Litscher *et al.* 2009). ZP2 y ZP3 se encuentran entre las proteínas más divergentes en mamíferos (Makalowski y Boguski 1998). Los estudios tanto inter como intraespecíficos han mostrado que ZP2 y ZP3 evolucionan adaptativamente, y que la selección positiva afecta a varios sitios localizados en la región de ZP3 implicada en el reconocimiento específico del espermatozoide (Swanson *et al.* 2001b; Jansa *et al.* 2003; Turner y Hoekstra 2006, 2008b).

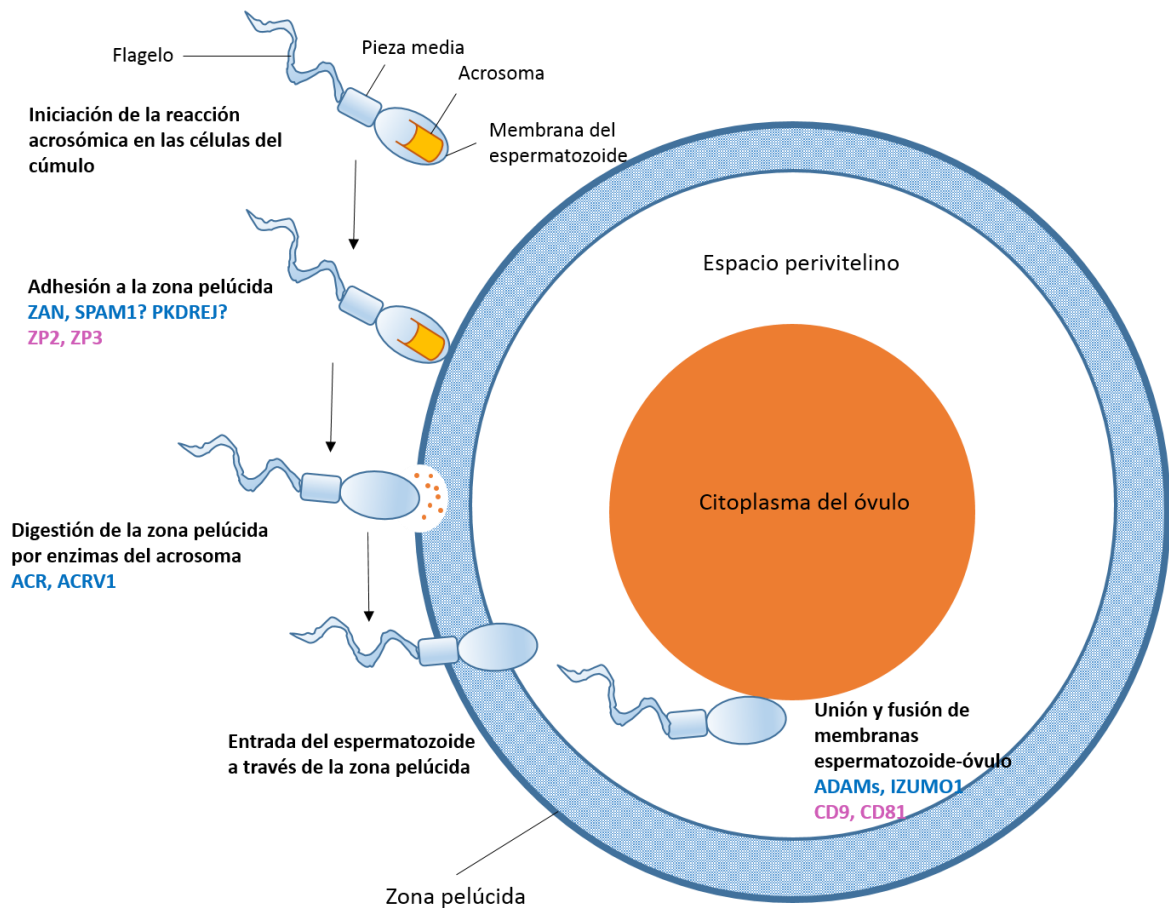


Figura 4. Esquema de las posibles interacciones entre el espermatozoide y el óvulo durante el proceso de fecundación. Las proteínas masculinas y femeninas con evidencia de selección positiva se indican en el azul y rosa respectivamente.

Aunque se ha identificado un gran número de proteínas en la membrana del espermatozoide, todavía no se han definido claramente qué proteínas se unen e interaccionan con las proteínas de la zona pelúcida (Ikawa *et al.* 2008; Avella *et al.* 2013). Se han descrito algunas proteínas candidatas implicadas en este proceso. La proteína SPAM1 (PH-20) es una hialuronidasa localizada en la membrana interna del acrosoma que se identificó inicialmente como una proteína que se unía a la zona pelúcida tras la reacción acrosómica (Primakoff *et al.* 1988). La zonadhesina (Zan), una proteína localizada en la membrana externa del acrosoma y que se expone a la superficie del espermatozoide durante la capacitación, está implicada en la especificidad de la adhesión espermatozoide-zona pelúcida en diversas especies (Tardif *et al.* 2010; Tardif y Cormier 2011). La β 1,4-galactosiltransferasa (GALT) es una proteína presente en la membrana del espermatozoide capaz de unirse a los oligosacáridos de ZP3 desencadenando la reacción acrosómica (Lopez *et al.* 1985; Macek *et al.* 1991; Rodeheffer y Shur 2002). Otra proteína que se ha propuesto como un receptor de la zona pelúcida es PKDREJ (Sutton *et al.* 2006) (Fig. 4), un factor implicado en el transporte y la regulación de la reacción acrosómica (Sutton *et al.* 2008).

Los estudios evolutivos sobre proteínas implicadas en fecundación han detectado que los genes *Spam1* (Torgerson *et al.* 2002; Swanson *et al.* 2003), y *Zan* (Swanson *et al.* 2003; Herlyn y Zischler 2005; Gasper y Swanson 2006) muestran evidencia de evolución adaptativa (Fig. 4), mientras que para *$\beta 4gal t1$* no se detectó señal de selección positiva (Swanson *et al.* 2003). Análisis tanto intra como interespecíficos en primates han revelado que la divergencia de PKDREJ es promovida por selección positiva (Hamm *et al.* 2007).

Durante la reacción acrosómica, las enzimas proteolíticas que digieren la zona pelúcida y permiten la penetración del espermatozoide se liberan del acrosoma. Una proteína a la que se le ha atribuido una función importante en este proceso es la acrosina (Acr) (Fig. 4), que es una abundante proteasa localizada en la matriz del acrosoma. No obstante, estudios en ratones deficientes en *Acr* sugieren que esta proteína no es esencial para tal proceso (Honda *et al.* 2002). Los análisis comparativos entre aves y mamíferos identificaron una señal de selección positiva en el gen *Acr* en ambos grupos (Berlin *et al.* 2008). Otra proteína contenida en el acrosoma que se cree implicada en la unión o la penetración de la zona pelúcida, *Acrv1* (Fig. 4), muestra una fuerte señal de selección positiva en roedores (Ramm *et al.* 2008).

Después de que el espermatozoide atraviese la zona pelúcida, el siguiente paso en la fecundación es la fusión entre las membranas de los gametos. Las proteínas implicadas en la fusión espermatozoide-óvulo también muestran alta variabilidad. En mamíferos, dos componentes esenciales para la fusión gamética son el receptor Cd9 presente en la membrana del óvulo (Miyado 2000) y la proteína Izumo1 en la superficie del espermatozoide (Inoue *et al.* 2005) (Fig. 4). Los análisis comparativos en mamíferos han revelado evidencia de evolución adaptativa en *Cd9* (Swanson *et al.* 2003) y en *Izumo1* (Dorus *et al.* 2010).

Una familia importante de proteínas implicadas en interacción espermatozoide-óvulo son las ADAMs. Las ADAMs son proteínas de membrana que presentan un dominio metaloproteinasas y un dominio desintegrina, por lo que van a tener funciones tanto proteolíticas como de adhesión celular (Primakoff y Myles 2000). De las 35 ADAMs que se han caracterizado hasta ahora, 18 son expresadas en tejidos reproductivos, y 9 de ellas se han localizado en la superficie del espermatozoide (Kim *et al.* 2006). Las ADAMs interaccionan con miembros de la familia integrina, y dado que algunas integrinas se expresan en el óvulo, estas proteínas pueden participar en la interacción gamética. (Primakoff y Myles 2000). En mamíferos, ADAM1, ADAM2 y ADAM3 son las proteínas mejor caracterizadas en la función reproductiva, y tienen un papel importante en la unión con la zona pelúcida, así como en la unión y fusión de membranas (Primakoff y Myles 2000; Evans 2002). Estudios moleculares han revelado que las ADAMs reproductivas evolucionan con mayor velocidad que las expresadas en tejidos somáticos y que sólo las ADAMs presentes en la membrana del espermatozoide muestran evidencia de selección positiva en los dominios de adhesión (Torgerson *et al.* 2002; Civetta 2003; Swanson *et al.* 2003; Glassey y Civetta 2004; Finn y Civetta 2010) (Fig. 4).

1.5. PROTEÓMICA DE LAS PROTEÍNAS REPRODUCTIVAS

Durante la última década se han desarrollado numerosos estudios de proteómica con el fin de analizar la evolución del proteoma de diversos tejidos reproductivos. Análisis del proteoma de las glándulas accesorias en roedores detectaron que las proteínas expresadas en la vesícula seminal muestran una mayor tasa de evolución (Dean *et al.* 2009). Además, se observó que el proteoma del fluido seminal de roedores presenta una gran diversificación tanto a nivel de composición como de proteínas (Ramm *et al.* 2009).

El proteoma del espermatozoide ha sido analizado en diversas especies de invertebrados (Dorus *et al.* 2006; Wasbrough *et al.* 2010) y mamíferos (Baker *et al.* 2007, 2008a,b; Peddinti *et al.* 2008; Dorus *et al.* 2010; Skerget *et al.* 2013). Algunos de estos trabajos se han centrado en analizar la evolución del proteoma del espermatozoide con el fin de identificar el impacto de las fuerzas selectivas sobre la función y la evolución del espermatozoide. Análisis en *Drosophila* mostraron que las proteínas integrales del espermatozoide, en contraste con las proteínas de las glándulas accesorias, se encuentran altamente conservadas, sugiriendo una compartimentación de la adaptación de las proteínas reproductivas (Dorus *et al.* 2006). Por otro lado, análisis proteómicos en ratón revelaron que las proteínas presentes en la membrana del espermatozoide evolucionan con mayor rapidez y muestran una señal más intensa de selección positiva (Dorus *et al.* 2010). Estos resultados sugieren que en el espermatozoide de mamíferos puede existir una compartimentación de la respuesta adaptativa a nivel subcelular.

1.6. CAUSAS DE LA EVOLUCIÓN ADAPTATIVA DE LAS PROTEÍNAS REPRODUCTIVAS

Tras observar que las proteínas relacionadas con la reproducción presentan una alta tasa de evolución en diversos grupos taxonómicos así como en diferentes etapas del proceso reproductivo, ha surgido la necesidad de conocer qué fuerzas son las causantes de esta divergencia adaptativa. Se han propuesto diversas fuerzas selectivas que podrían influir sobre la evolución de las proteínas reproductivas, de modo que diferentes fuerzas pueden actuar dependiendo del organismo y de la función de la proteína. En muchos casos, es posible que varias fuerzas estén implicadas en la evolución de un mismo gen. A continuación se resumen las diferentes fuerzas selectivas que se han propuesto para explicar la evolución adaptativa de las proteínas reproductivas.

1.6.1. Resistencia a patógenos

El tracto reproductivo femenino presenta un sistema inmune capaz de responder ante agentes infecciosos (Wira *et al.* 2005). Por tanto, cuando el espermatozoide penetra en el tracto reproductivo de la hembra, algunas de las moléculas transferidas pueden ser reconocidas como agentes patógenos, desencadenando una respuesta inmune que destruye a los espermatozoides. Debido a esto, el

espermatozoide deberá contener proteínas capaces de modular la respuesta inmune de la hembra (Harris *et al.* 2006). Se han identificado multitud de proteínas tanto en el fluido seminal como en la superficie del espermatozoide implicadas en respuesta inmune (Pilch y Mann 2006; Dorus *et al.* 2012). Un estudio reciente ha revelado que una alta proporción de proteínas detectadas bajo selección positiva en el proteoma del espermatozoide son del sistema inmune (Dorus *et al.* 2010). Es posible que proteínas inmunes del tracto reproductivo de la hembra y del espermatozoide estén coevolucionando adaptativamente como resultado de una “carrera de armamentos” entre el hospedador y el patógeno (Good y Nachman 2005; Dean *et al.* 2009; Dorus *et al.* 2010). No obstante, se ha propuesto recientemente que las proteínas inmunes del espermatozoide podrían tener funciones pleiotrópicas tanto en la respuesta inmune de la célula como en la reproducción, por lo que tanto fuerzas inmunológicas como sexuales podrían controlar conjuntamente la evolución de estas proteínas (Dorus *et al.* 2012).

1.6.2. Selección críptica de la hembra

La selección críptica, en especies con fecundación externa, describe el fenómeno por el que un óvulo tiene preferencia por unirse a un espermatozoide que presenta un genotipo determinado para algún receptor de la membrana, teniendo menor afinidad por espermatozoides que presentan alelos diferentes (Eberhard 1996). Un ejemplo de selección críptica en invertebrados marinos es el que se estudió en poblaciones del erizo de mar *Echinometra*, donde los óvulos muestran preferencia por unirse a espermatozoides con el mismo alelo de bindina (Levitan y Ferrell 2006).

En el caso de las especies con fecundación interna, también se ha observado selección críptica dentro del tracto reproductivo de la hembra en grupos diversos como insectos, reptiles y aves (Ward 2000; Wagner *et al.* 2004; Moore *et al.* 2009). El mecanismo de selección críptica en estas especies está basado en la capacidad de las hembras para almacenar y seleccionar espermatozoides dentro de los órganos de almacenaje.

1.6.3. Competición espermática

Como se mencionó anteriormente, la competición espermática surge cuando los espermatozoides (fecundación externa) o eyaculados (fecundación interna) de múltiples machos compiten por fecundar los óvulos de una misma hembra (Parker 1970). En esta situación, la competición espermática va a promover la adaptación de diferentes rasgos de los espermatozoides y el eyaculado con el fin de incrementar el éxito reproductivo de los machos. La competición espermática puede ejercer una presión selectiva sobre la cantidad de esperma que se va a inseminar, la viabilidad, la morfología y la velocidad del espermatozoide (Pizarri y Parker 2009). Además, los espermatozoides pueden adaptarse para ser los mejores en interactuar con los componentes de la zona pelúcida, así como en unirse y fusionarse con la membrana del óvulo (Swanson y Vacquier 2002). Se ha asumido que la adaptación de rasgos fenotípicos del espermatozoide en términos de

competición espermática es el resultado de la evolución adaptativa de una serie de proteínas reproductivas que regulan esos rasgos (Swanson and Vacquier 2002). Por lo tanto, la competición espermática se ha sugerido como una causa potencial de la rápida evolución de las proteínas reproductivas.

1.6.4. Conflicto sexual

El conflicto sexual ocurre cuando los intereses de los dos sexos difieren, y los rasgos que son favorecidos para un sexo son perjudiciales para el otro (Chapman *et al.* 2003). El resultado del conflicto sexual va a ser un proceso de coevolución entre el macho y la hembra.

En el caso de mamíferos, la competición espermática va a desembocar en situaciones de conflicto sexual, ya que cuando hay un exceso de espermatozoides en el tracto reproductivo de la hembra, va a existir un mayor riesgo de polispermia. La polispermia en la mayoría de los casos supone una pérdida del embrión, por lo que es perjudicial para la hembra. Ante esta situación, la hembra puede responder reduciendo la receptividad del óvulo a los espermatozoides para evitar la polispermia. Esto va a promover que los espermatozoides tengan que experimentar nuevos cambios para volver a maximizar la afinidad por el óvulo, resultando así un proceso de coevolución entre los sexos (Swanson y Vacquier 2002; Clark y Dell 2006).

Por lo tanto, el conflicto sexual puede influir sobre proteínas de la superficie del espermatozoide y del óvulo, así como sobre proteínas del fluido seminal y del tracto reproductivo de la hembra, siendo una fuerza importante en la evolución de proteínas mediando interacciones post-cópula entre el macho y la hembra.

1.7. COMPETICIÓN ESPERMÁTICA Y LA EVOLUCIÓN DE PROTEÍNAS REPRODUCTIVAS

Durante los últimos años han sido numerosos los estudios que han intentado demostrar la hipótesis de que la tasa evolutiva de las proteínas reproductivas, así como los niveles de selección positiva, deben incrementar con la intensidad de la competición espermática entre las especies. No obstante, relativamente pocos estudios han encontrado evidencia de que los genes reproductivos evolucionan bajo la influencia de la competición espermática.

Hasta ahora se han utilizado dos métodos para tratar de analizar el efecto de la competición espermática sobre los genes reproductivos. El primer método consiste en usar medidas continuas de los niveles de competición espermática. La variable que se ha utilizado con mayor frecuencia en estos análisis es el tamaño relativo de los testículos, ya que el tamaño testicular de cada especie es un buen indicativo de su intensidad de competición espermática (Gomendio *et al.* 1998). En este método, se estiman valores promedio de ω para cada especie de la filogenia, (Yang 1998; Yang y Nielsen 1998). Los valores específicos de ω entonces se relacionan con la variable continua de

competición espermática en un análisis de regresión, donde una correlación significativa se interpreta como una evidencia de que la competición espermática está dirigiendo la evolución del gen reproductivo en cuestión. El primer estudio que demostró un efecto de la competición espermática sobre la evolución de un gen reproductivo utilizando este método fue el de Dorus *et al.* (2004), quienes encontraron que la tasa evolutiva del gen *SEMG2* se correlaciona positivamente con los niveles de promiscuidad femenina y el tamaño relativo de testículo en primates (Fig. 5). Otro estudio consiguió establecer una asociación entre la evolución de la proteína zonadhesina (ZAN) y niveles de competición espermática en primates (Herlyn y Zischler 2007). En este caso, se utilizó una medida de dimorfismo sexual en tamaño como índice de competencia espermática, detectando una correlación negativa, ya que un menor dimorfismo sexual indica mayor intensidad de competición espermática.

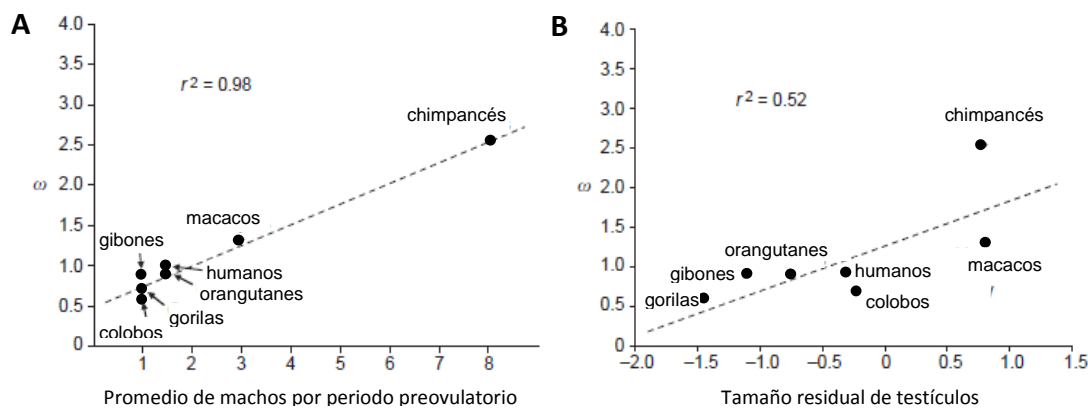


Figura 5. Relación entre la tasa evolutiva de *SEMG2* y competición espermática en primates. A. Correlación entre niveles de promiscuidad, medidos como el promedio de machos apareando con cada hembra durante el periodo preovulatorio, y los valores específicos de ω (dN/dS); B. Correlación entre el tamaño residual de testículos corregido por el tamaño corporal y los valores específicos de ω . Modificado de Dorus *et al.* (2004),

El segundo método es categórico y consiste en clasificar las especies analizadas en función de sus sistemas de apareamiento. Dentro de una filogenia, las ramas de cada linaje pueden marcarse de forma diferente en función de si presentan altos o bajos niveles de competición espermática. Esta filogenia marcada puede entonces ser usada para comparar diferentes modelos de evolución. En caso de que los modelos de selección positiva se ajusten mejor en las ramas de los linajes de alta competición espermática que en las de los linajes de baja competición espermática, esto se interpreta como evidencia de una asociación entre competición espermática y selección positiva para el gen en estudio (Zhang *et al.* 2005). Este método fue utilizado por Ramm *et al.* (2008) para analizar la evolución de 7 genes de roedores. A pesar de que se detectó evidencia de evolución adaptativa en 5 genes, solo se encontró señal de selección positiva concentrada en los linajes de alta competición espermática para el gen *Svs2*, que es homólogo al gen *SEMG2* de primates. Estos análisis fueron

también aplicados a los genes *SEMG1* y *SEMG2* de primates, detectando selección positiva asociada a competición espermática en *SEMG2*, corroborando por tanto los resultados de Dorus *et al.* (2004) y la concordancia de ambos métodos. Posteriormente, Finn y Civetta (2010) usaron este método para analizar la evolución de las ADAMs en primates. Se detectó evidencia de evolución adaptativa en 12 de las 13 ADAMs expresadas en el macho. No obstante, solo 3 genes (*ADAM2*, *ADAM18* y *ADAM23*) presentaron señal de selección positiva en los linajes más poliándricos.

Independientemente de los métodos y las especies utilizadas, son pocos los genes que se han encontrado bajo el efecto de la competición espermática, a pesar de que son numerosos los genes reproductivos que han mostrado evidencia de selección positiva.

1.8. LOS ROEDORES COMO MODELO EN ESTUDIOS DE COMPETICIÓN ESPERMÁTICA

Los roedores (orden Rodentia) componen el orden más abundante y más diverso de los mamíferos. Hasta el momento se han identificado unas 2.300 especies de roedores, abarcando aproximadamente el 40% de todas las especies de mamíferos (Wilson y Reeder 2005). Su distribución geográfica es inmensa, encontrándose en todos los continentes menos en la Antártida. La superfamilia de roedores más abundante es la superfamilia Muroidea, que contiene casi un tercio de las especies de mamíferos, e incluyen las familias Muridae y Cricetidae, que presentan una gran diversidad taxonómica y una amplia distribución geográfica (Steppan *et al.* 2004)

Los roedores son un grupo ideal para el desarrollo de estudios de competición espermática por diversas razones. Por un lado, los roedores presentan una gran diversidad de sistemas de apareamiento, por lo que cubren prácticamente todo el rango de niveles de competición espermática (Kenagy y Trombulak 1986; Tourmente *et al.* 2011b). Por otro lado, se ha descubierto un gran número de adaptaciones a la selección sexual post-cópula en roedores y a diferentes niveles. Se ha observado que la competición espermática influye sobre el comportamiento reproductivo en roedores, aumentando la frecuencia de cópulas (delBarco-Trillo y Ferkin 2004; Stockley y Preston 2004). Ramm *et al.* (2005) detectaron una correlación entre el tamaño relativo de testículos y el tamaño de las glándulas accesorias, sugiriendo que el sistema reproductivo de los roedores presenta una diversidad anatómica asociada a competición espermática. A nivel celular, los roedores muestran una enorme variedad en las dimensiones y morfología del espermatozoide (Roldan *et al.* 1992; Gomendio y Roldan 2008; Tourmente *et al.* 2011b; Varea Sánchez *et al.* 2013). Una amplia variedad también se ha observado en rasgos fisiológicos del espermatozoide como la viabilidad (Gómez Montoto *et al.* 2011a), velocidad de natación (Gomendio y Roldan 2008; Gómez Montoto *et al.* 2011b), o niveles de ATP (Tourmente *et al.* 2013). La adaptación de rasgos reproductivos a la competición espermática se ha podido observar incluso en grupos de especies muy próximas, como

es el caso del género *Mus*. Un estudio en 4 especies de este grupo reveló que la competición espermática favorece una mayor proporción de espermatozoides preparados para fecundar y para responder a las señales del óvulo (Gomendio *et al.* 2006). En un análisis posterior con 3 de estas 4 especies, se mostró que la competición espermática es capaz de incrementar tanto la competitividad del esperma como la defensividad del óvulo, promoviendo barreras asimétricas de aislamiento reproductivo (Martin-Coello *et al.* 2009).

La observación de que especies muy próximas presentan notables diferencias fenotípicas hace que los roedores también supongan un buen modelo para análisis de evolución molecular. Los estudios de especies entre las que se ha producido un aislamiento reproductivo reciente son importantes para identificar los primeros cambios que generan diferencias entre especies, y comprender el significado adaptativo de tales cambios. Además, la comparación de secuencias no muy divergentes incrementa el poder de los análisis de selección positiva (Anisimova *et al.* 2001). Así, son múltiples los genes reproductivos que han mostrado evidencia de selección positiva cuando se han comparado especies de roedores estrechamente relacionadas (Jansa *et al.* 2003; Podlaha *et al.* 2005; Turner y Hoekstra 2006; Ramm *et al.* 2008; Lüke *et al.* 2011).

2. OBJETIVOS

El objetivo principal de esta Tesis Doctoral fue el de realizar un estudio a diferentes niveles sobre la evolución de las proteínas del espermatozoide en roedores, con el fin de identificar proteínas que experimentan una evolución adaptativa en respuesta a las fuerzas de selección sexual, y más concretamente la competición espermática.

Los diferentes estudios que se presentan en esta Tesis Doctoral se desarrollaron con los siguientes objetivos:

- Determinar si las proteínas que participan en diferentes eventos de la vida del espermatozoide presentan diferencias en la tasa evolutiva. En tal caso, identificar qué procesos están sujetos a una mayor intensidad de selección positiva.
- Estudiar la evolución de una proteína implicada en la motilidad del espermatozoide (Catsper1) y observar si la divergencia de su secuencia, tanto a nivel molecular como estructural, se relaciona con los niveles de competición espermática, así como con la velocidad de natación del espermatozoide.
- Analizar la evolución molecular de una proteína implicada en el transporte del espermatozoide y en la regulación de la reacción acrosómica (PKDREJ) y determinar si su tasa evolutiva está asociada con la competición espermática.
- Aplicar análisis de coevolución sobre dos proteínas implicadas en la fusión espermatozoide-óvulo (Izumo1 y Cd9) con el fin de predecir si estas proteínas presentan una interacción directa.
- Desarrollar un estudio de proteómica comparativa entre tres especies de roedores con diferentes niveles de competición espermática con el fin de detectar cambios de abundancia en proteínas del espermatozoide asociados con esta fuerza selectiva.

3. MATERIAL Y MÉTODOS

3.1. ANIMALES

3.1.1. Especies utilizadas

Para los diversos estudios se emplearon diferentes especies de roedores pertenecientes a la superfamilia Muroidea. Concretamente se trabajó con especies de las familias Muridae y Cricetidae. Todas las especies de la familia Muridae utilizadas pertenecen a la subfamilia Murinae, mientras que las especies de la familia Cricetidae pertenecen a dos subfamilias: Arvicolinae y Cricetinae (Tabla I). Las relaciones filogenéticas de las especies se muestran en la Fig. 6.

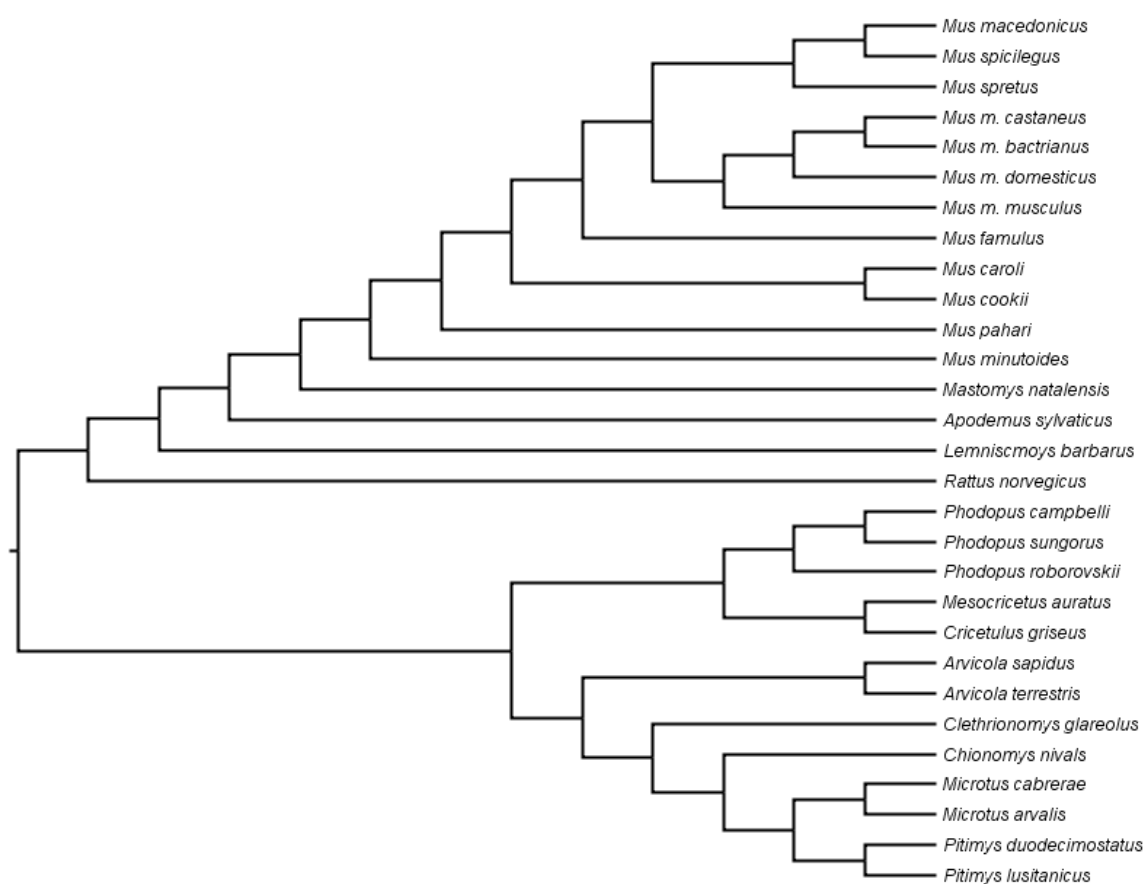


Figura 6. Cladograma mostrando las relaciones filogenéticas de las especies de muroideos. Basado en las filogenias de Suzuki *et al.* 2004; Steppan *et al.* 2005; Martín-Coello *et al.* 2009; Fabre *et al.* 2012

Tabla I. Especies de roedores muroideos estudiadas y sus localidades de origen.

Familia	Subfamilia	Género	Especie	Subespecie	Localidad de captura
Muridae	Murinae	<i>Apodemus</i>	<i>sylvaticus</i>		Cerrato, Castilla-León, España
		<i>Lemniscomys</i>	<i>barbarus</i>		Túnez
		<i>Mastomys</i>	<i>natalensis</i>		Durban, Sudáfrica
		<i>Mus</i>	<i>caroli</i>		Indonesia
			<i>cookii</i>		Tailandia
			<i>famulus</i>		Kotagiri, India
			<i>macedonicus</i>		Slantchev Briag, Bulgaria
			<i>minutoides</i>		Ciudad del Cabo, Sudáfrica
			<i>musculus</i>	<i>musculus</i>	Alazani, Georgia
				<i>castaneus</i>	Masinagudi, India
				<i>domesticus</i>	Azzemour, Marruecos
				<i>bactrianus</i>	Machad, Irán
			<i>spicilegus</i>		Ucrania
			<i>spretus</i>		Azzemour, Marruecos
			<i>pahari</i>		Tailandia
		<i>Rattus</i>	<i>norvegicus</i>		
Cricetidae	Arvicolinae	<i>Arvicola</i>	<i>sapidus</i>		Soria, Castilla-León, España
			<i>terrestris</i>		Palencia, Castilla-León, España
		<i>Clethrionomys</i> (*)	<i>glareolus</i>		Montseny, Barcelona, España
		<i>Chionomys</i>	<i>nivalis</i>		Navacerrada, Madrid, España
		<i>Microtus</i>	<i>arvalis</i>		Castilla-León, España
			<i>cabrerae</i>		Cahapinería, Madrid, España
		<i>Pitimys</i> (**)	<i>duodecimostatus</i>		Castilla-León, España
			<i>lusitanicus</i>		El Ventorrillo, Madrid, España
	Cricetinae	<i>Cricetulus</i>	<i>griseus</i>		Kumtagh, China
		<i>Mesocricetus</i>	<i>auratus</i>		Alepo, Siria
		<i>Phodopus</i>	<i>campbelli</i>		Altai, Rusia
			<i>roborovskii</i>		Tuva, Rusia
			<i>sungorus</i>		Kazajstán

(*) Especie comúnmente denominada *Myodes glareolus*.

(**) Especies actualmente reconocidas dentro del género *Microtus*.

Los individuos de las especies *Mus caroli*, *Mus cookii*, *Mus famulus*, *Mus macedonicus*, *Mus m. musculus*, *Mus m. castaneus*, *Mus m. domesticus*, *Mus m. bactrianus*, *Mus minutoides*, *Mus spicilegus*, *Mus spretus*, *Mus pahari*, *Lemniscomys barbarus*, *Mastomys natalensis*, *Cricetulus griseus*, *Mesocricetus auratus*, *Phodopus campbelli*, *Phodopus roborovskii* y *Phodopus sungorus* se obtuvieron de colonias establecidas y procedentes de poblaciones silvestres, las cuales se mantuvieron en cautividad durante unas pocas generaciones. En el caso de las especies de *Mus*, los individuos fueron suministrados por el Institut des Sciences de l'Évolution, CNRS-Université Montpellier 2 (Francia). El resto de los individuos fueron adquiridos de diversos centros de cría en

España. Los individuos de las especies *Apodemus sylvaticus*, *Arvicola sapidus*, *Arvicola terrestris*, *Clethrionomys glareolus*, *Chionomys nivalis*, *Microtus arvalis*, *Microtus cabreræ*, *Pitimys duodecimostatus* y *Pitimys lusitanicus* fueron capturados en el campo durante la época de apareamiento.

Dado que no se dispuso de ejemplares de *Rattus norvegicus*, todos sus datos necesarios para los análisis fueron obtenidos de bases de datos de acceso público o de la literatura.

3.1.2. *Mantenimiento de los animales*

Los animales fueron alojados en cajas individuales y mantenidos en un animalario bajo condiciones estándar, con una temperatura controlada de 20-24°C, un fotoperiodo de 14 horas de luz y 10 horas de oscuridad, y disponiendo de agua y alimento *ad libitum*.

El cuidado y tratado de los animales siguió la normativa RD1201/2005 de Regulación de la Protección de la Fauna Española, aprobada conforme a la ley 2003/65 de Regulación de la Unión Europea para el Cuidado y Uso de los Animales de Laboratorio. Los estudios que requirieron del uso de estos animales se aprobaron por el Comité de Bioética del Consejo Superior de Investigaciones Científicas (CSIC, Spain).

3.2. OBTENCIÓN DE MUESTRAS

3.2.1. *Sacrificio y disección del animal*

Los animales fueron sacrificados por la técnica de dislocación cervical y se pesaron en una balanza. Se tomaron medidas del tamaño corporal y se procedió a su disección a la altura de la línea media ventral dejando expuesto el aparato reproductor.

En el caso de los machos, se extrajeron los testículos y los epidídimos junto con los conductos deferentes. El contenido del conducto deferente se llevó hacia el epidídimo utilizando una pinza, y a continuación se eliminó el conducto deferente cortándolo con una tijera. Los epidídimos se colocaron en una placa Petri con 1-3 ml de medio Tyrode modificado con Hepes (mT-H) precalentado a 37°C (Shi y Roldan 1995). A continuación se hicieron incisiones en el epidídimo y se dejaron en el medio durante 10 minutos para permitir la salida de los espermatozoides maduros. Transcurrido este tiempo, se recogieron los epidídimos y se almacenaron a -80°C. De la suspensión espermática se tomaron muestras que se utilizaron para analizar diversos parámetros. En el caso de las hembras, se extrajeron los ovarios tras separarlos de los oviductos.

Tanto los testículos como los ovarios fueron depositados en nitrógeno líquido tras su extracción y posteriormente fueron almacenados a -80°C hasta su uso. De todos los individuos además se tomaron muestras de vesícula seminal, hígado, bazo, cola y músculo y se almacenaron a -80°C. Los cadáveres también fueron almacenados a -80°C.

3.3. ANÁLISIS DE BIOLOGÍA MOLECULAR

3.3.1. *Extracción de ácidos nucleicos*

Para la amplificación de algunos genes se procedió a la obtención del ADN genómico de las especies de roedores analizadas. Las muestras de partida fueron fragmentos de cola o de pata conservados a -80°C, ya que estos tejidos presentan baja actividad transcripcional y por tanto poca cantidad de ARN. El procedimiento se llevó a cabo con el kit comercial E.Z.N.A® Tissue DNA Kit (Omega, Madrid, España) siguiendo las instrucciones del proveedor. Se obtuvo DNA genómico de un mínimo de cinco individuos por especie, excepto en las especies *Mus m. bactrianus*, *Mus m. castaneus*, *Mus macedonicus* y *Mus famulus*, de las que se disponían muestras de tres individuos.

Para algunas especies de la familia Muridae se obtuvo el ARN total de testículo y de ovario utilizando el kit comercial total RNA kit I (Omega). En el caso de los testículos, se partió de una muestra inicial de ~50 mg que se homogenizó en 700 µl de buffer TRK, lo que implicó el uso de un fragmento o de todo el tejido en función de la especie. De los ovarios se utilizó todo el tejido, y al tener un peso de ~20 mg se homogeneizaron en 300 µl de buffer TRK. La obtención de ARN total se llevó a cabo para un mínimo de tres individuos por especie.

La síntesis *in vitro* de ADN complementario (ADNc) se realizó con el kit comercial SuperScript III First-Strand Synthesis System (Life technologies, Madrid, España), según el protocolo descrito por el fabricante. El procedimiento consistió en la desnaturalización del ARN molde, empleando una cantidad de ARN total de partida de 1-2 µg, durante 5 min a 65°C y en presencia de oligo (dT)₂₀ y dNTPs. La muestra se pasó rápidamente a hielo y se añadió Syntesis buffer, DTT (ditiotreitól), RNaseOUT (inhibidor de las RNasas), MgCl₂ y la enzima Superscript III Reverse Transcriptase 50 U (Life technologies) a las concentraciones recomendadas por el fabricante. La síntesis de ADNc tuvo lugar a 42°C durante 50 min, y entonces la muestra se incubó a 70°C durante 15 min para desnaturalizar a la enzima. Para hidrolizar el ARN molde se aplicó 1 µl de la enzima RNasa H suministrada por el kit y se incubó a 37°C durante 20 minutos. Las amplificaciones de ADNc se obtuvieron de tres individuos por especie.

Tras finalizar las extracciones de ADN, ARN y ADNc, se evaluó la pureza y la concentración total de los ácidos nucleicos respectivos mediante el espectrofotómetro NanoDrop 1000 (Thermo scientific), utilizando el buffer de elución respectivo como control antes de la medición.

3.3.2. *Amplificación de secuencias por PCR*

Los fragmentos de interés fueron amplificados por la técnica de reacción en cadena de la polimerasa (PCR) partiendo bien de ADN genómico (PCR convencional) o de ADN complementario (ADNc) obtenido por retrotranscripción (RT-PCR). En cada caso, se diseñaron cebadores para

amplificar los genes de interés de manera específica. Los cebadores específicos se diseñaron en base a secuencias flanqueando la región a amplificar que se encuentran conservadas entre especies cercanas. En el caso de roedores, las especies que a menudo se utilizaron como referencia en el diseño de los cebadores fueron *M. musculus* y *Rattus norvegicus* para la familia Muridae, y *Microtus ochrogaster* y *Mesocricetus auratus* cuando se trabajó con especies de la familia Cricetidae. Los genomas de estas especies se encuentran disponibles en la base de datos del NCBI (<ftp://ftp.ncbi.nlm.nih.gov/genomes>).

Los cebadores que se utilizaron para amplificar las regiones de los diferentes genes que se estudiaron se representan en la tabla II.

Las reacciones de PCR se prepararon en un volumen de 50 µl conteniendo PCR Gold Buffer 1x (Roche Applied Science, Barcelona, España), 2,5mM de MgCl₂ (Roche Applied Science), 0,8 mM de dNTPs (Applied Biosystems, Barcelona, España), 0,25 mM de cada cebador (sintetizados por Life Technologies), 2 unidades de polimerasa AmpliTaq Gold (Roche Applied Science) y 20-200 ng/µl de ADN genómico o ADNc molde. Las reacciones de amplificación se llevaron a cabo en un termociclador Veriti (Applied Biosystems).

El programa de amplificación consistió en un paso de activación de la polimerasa de 5 min a 95°C, seguido de 30-40 ciclos compuestos de 3 pasos:

- 1) Desnaturalización a 95°C durante 40 segundos;
- 2) Acoplamiento de los cebadores al molde. 1 min a la temperatura correspondiente al T_m de los cebadores. Cuando las T_m de los cebadores fueron diferentes, se aplicó un gradiente de temperatura en este paso.
- 3) Elongación de la hebra a 72°C durante un tiempo variable en función del fragmento a amplificar. Normalmente se aplicó 1 min por cada 1000 bases (Kb) de secuencia.

Tras completarse todos los ciclos, se aplicó un paso de elongación final de 8 min a 72°C. Una vez finalizado el programa, las reacciones se almacenaron a -20°C.

Tabla II. Cebadores utilizados para la amplificación de los genes de estudio

Gen	Especies de referencia	Cebador	Secuencia	Tm (°C)
<i>Catsper1</i>	<i>M. musculus</i> (NM_139301), <i>Rattus norvegicus</i> (XM_001070492)	Csp1 Fw1	GAGCTSTCTTTGTTCCAGGG	62
		Csp1 Mus Fw1	ATGGATCAATCTTCAAGGAG	64
		Csp1 Mus Fw3	CATGGATCAAYCTTCAAGGA	62
		Csp1 Rv1	GRACCCAGGAAAGACTCAC	64
		Csp1 Rv2	AGGAAAGACTCACTTTTGCG	62
		Csp1 Mus Rv2	TCACTCARAGAGGCYGATTC	64
<i>Pkdrej</i>	<i>M. musculus</i> (NM_011105), <i>Rattus norvegicus</i> (NM_001134866), <i>Microtus ochrogaster</i> (GCA_000317375.1), <i>Mesocricetus auratus</i> (GCA_00349665.1)	Pkdrej1-Fw	TTACGGGGTGTATGTGTTTAAC	62
		Pkdrej voles Fw1	CTGTCCATCAACAGGTGGGA	62
		Pkdrej Ham Fw2	CTACATTTGGGTAAAAAAC	52
		Pkdrej1-Rv	CTGGGTAAATTTGGTAATCAC	58
		Pkdrej3-Rv	TCGTCCTCTGTGTTTCATAGCAA	64
		Pkdrej voles Rv1	AGTACCCTGGACGAGATCTG	62
		Pkdrej Ham Rv2	AGATAGGCTCTATTATTTTG	52
<i>Izumo1</i>	<i>M. musculus</i> (NM_001018013), <i>Rattus norvegicus</i> (NM_001017514)	Izumo1 Fw1	GTTTGGGGTGAGAGRATCCTC	60
		Izumo1 Fw3	GACTGATTTTGGCTTCTGTTTA	60
		Izumo1 Fw4	ATGGGGCCGCATTTTACACT	60
		Izumo1 Rv1	CCAGAAGTCAGAACCAGTCC	62
		Izumo1 Rv4	GTTTTCTGTTGCCTCGCTCT	60
<i>Cd9</i>	<i>M. musculus</i> (NM_007657), <i>Rattus norvegicus</i> (NM_053018)	Cd9 Fw1	TTGTGCGAGTCCCTTCTGTCC	60
		Cd9 Fw2	CAGTCGTTCTGTCCTCTTGT	60
		Cd9 Rv1	AAATGGCCCGGACAGTCTTC	60
		Cd9 Rv3	AGAGTCTGCCCAACCCCGA	62

3.3.3. Electroforesis en gel de agarosa

Con el fin de detectar los fragmentos amplificados tras la PCR, se tomaron 3 µl del producto de la PCR y se mezclaron con tampón azul de carga (6X) antes de ser aplicados a un gel de agarosa (Hispanlab, Madrid, España) de concentración comprendida entre el 1% y el 2% y al que fue añadido un 1% de SYBR® Safe diluido (Life technologies) para el marcaje de las bandas. La electroforesis se llevó a cabo en una cubeta con tampón TBE 1x y a una potencia comprendida entre los 90V y los 110V, dependiendo de las dimensiones del gel. Junto con las muestras se cargaron marcadores de peso molecular de 100 pb y de 1 Kb (Biotools, Madrid, España) para comprobar el tamaño y la concentración de los amplicones.

Una vez finalizada la electroforesis, los geles fueron colocados en un transiluminador de luz UV y visualizados con el programa GelImage (Bio-Rad, Madrid, España).

3.3.4. Purificación de los fragmentos amplificados

Los productos de PCR que resultaron en la amplificación del fragmento de interés fueron purificados usando el kit comercial E.Z.N.A® Cycle Pure Kit (Omega) y utilizando el protocolo de lavado por centrifugación facilitado por el proveedor.

Cuando más de un fragmento aparte de nuestra región de interés fue amplificado de manera inespecífica, todo el producto de PCR fue cargado en un gel de agarosa al 0,8% y la banda de interés fue extraída y purificada utilizando el kit comercial E.Z.N.A® Gel Extraction Kit (Omega) siguiendo el protocolo por centrifugación.

3.3.5. Obtención y procesamiento de las secuencias

Los productos amplificados fueron introducidos en un equipo de secuenciación automática de ADN (Secugen S.L., Madrid, España) y las dos hebras fueron secuenciadas. Para la secuenciación completa de los fragmentos se necesitó diseñar una pareja de cebadores internos para cada fragmento (Tabla III). Estos cebadores hacen que la polimerasa trabaje en dirección a los extremos de las hebras para poder secuenciarlos, y deben tener una región de solapamiento para no perder información.

Las secuencias obtenidas se analizaron para determinar su homología usando el algoritmo blastn del programa BLAST (<http://www.ncbi.nlm.nih.gov/blast/>). A continuación, las secuencias fueron procesadas y analizadas por el software BioEdit (Hall, 1999). Aquellas posiciones que no se secuenciaron correctamente fueron corregidas basándonos en la lectura del cromatograma. Los sitios heterocigotos fueron designados como sitios ambiguos basados en la nomenclatura de la IUPAC (<http://www.bioinformatics.org/sms/iupac.html>). Las secuencias de nucleótidos se delimitaron a la región codificante y se tradujeron *in silico* a secuencias de aminoácidos usando la herramienta Translate de Bioedit. De este modo se comprobó si las secuencias presentaban errores que implicaban cambios en el marco de lectura. Los alineamientos de múltiples secuencias se llevaron a cabo usando el programa ClustalW (Thompson *et al.* 1994) (<http://www.clustal.org/>), variando los diferentes parámetros de penalizaciones para evaluar la robustez de los alineamientos. Las secuencias de individuos pertenecientes a la misma especie fueron alineadas y con BioEdit se creó una secuencia consenso.

Tabla III. Cebadores internos utilizados para secuenciar los fragmentos de estudio.

Gen	Cebador	Secuencia
<i>Catsper1</i>	Csp1 int Fw1	AGAGTGGACCTCGTATTGAC
	Csp1 int Fw2	CATAGACCAGTGAACCTTTG
	Csp1 int Fw3	CCCAWGRGTGGATCTCATCCC
	Csp1 int Rv1	ACTGGAGCGATAGGAAATGG
	Csp1 int Rv2	CTGTTCTGCTGGGGGTAAAG
	Csp1 int Rv3	GTGGGAAGCAYCCTTTYTGC
<i>Pkdrej</i>	PKDREJ int Fw1	ATGTTTGAAGTGGGCAGGGCAAG
	PKDREJ int Fw3	ATGTCTGTCACCCAGAACAG
	PKDREJ int Rv2	AAATTCCTAAAGGCAAATG
	PKDREJ int Rv3	AATTTTCAATGCATGAAATG
<i>Izumo1</i>	Izumo1 int Fw1	CCACATAGAGGTACATCGCT
	Izumo1 int Rv1	CCTTGGTTGATGGTATCTAG
<i>Cd9</i>	Cd9 int Fw1	CTGCTGTGGAGCTGTACAAG
	Cd9 int Fw2	GGAGTGTACATTCTGATTGGAG
	Cd9 int Rv1	CTCCTGGAGTTCTTTAATCAC
	Cd9 int Rv2	AACTTTTGGTAGGTGTCCTTG

3.4. ANÁLISIS FILOGENÉTICOS

3.4.1. Selección del modelo de evolución

No todas las secuencias de ADN evolucionan de la misma forma. Por esta razón, se han descrito numerosos modelos de evolución que asumen diferentes tasas de sustitución de nucleótidos entre las secuencias. Para encontrar el modelo que mejor describe la evolución de nuestras secuencias, se empleó el programa JModelTest (Posada 2008). Este programa compara un total de 56 modelos de evolución utilizando diferentes algoritmos como Máxima Verosimilitud o el Criterio de Información Akaike (AIC) y selecciona el de mayor probabilidad.

3.4.2. Construcción de árboles filogenéticos

Se construyeron árboles filogenéticos de roedores a partir de los genes analizados, con el fin de observar si la evolución de estos genes sigue la misma tendencia evolutiva de las especies, o si por el contrario está dirigida por otras fuerzas evolutivas. Se utilizaron dos tipos de métodos para la construcción de filogenias:

- Máxima verosimilitud (*Maximum Likelihood*): método estadístico que consiste en la atribución de distintas probabilidades a las diferentes topologías que puede presentar un árbol filogenético, basándose en un modelo de evolución. Los análisis fueron llevados a cabo con el programa PhyML 3.0 (Guindon *et al.* 2010)

- Inferencia Bayesiana: algoritmo basado en estadística bayesiana que trata de encontrar el árbol que mejor agrupa las especies mediante la estimación de una probabilidad posterior en cada clado. Los análisis bayesianos se realizaron con el programa MrBayes 3.1 (Ronquist y Huelsenbeck 2003).

3.5. ANÁLISIS EVOLUTIVOS

La estimación de la tasa evolutiva de las secuencias codificantes a menudo se lleva a cabo mediante el cálculo del valor ω , que equivale al cociente entre el número de mutaciones no sinónimas por sitios no sinónimos (d_N) y el número de mutaciones sinónimas por sitios sinónimos (d_S). El valor de ω nos da información sobre el tipo de evolución que experimenta la secuencia analizada. Así, un valor de $\omega < 1$ indica que la secuencia se encuentra conservada y que evoluciona bajo selección purificante; un valor de ω igual o cercano a 1 indica que la secuencia está sujeta a relajación funcional y que por tanto sufre cambios por evolución neutral; por último, un valor de $\omega > 1$ indica que la secuencia sufre mutaciones no sinónimas fijadas por selección positiva, sugiriendo que estos cambios son adaptativos.

Los análisis evolutivos fueron llevados a cabo utilizando la herramienta Codeml incluida en el programa PAML (Yang 2007). Codeml implementa el modelo de sustitución de codones (trinucleótidos que codifican para un aminoácido) de Goldman y Yang (1994) para estimar tasas evolutivas en secuencias de ADN que codifican para proteínas.

3.5.1. *Estimación de tasas evolutivas específicas para cada codon*

El modelo de Goldman y Yang (1994) permite estimar una tasa evolutiva promedio a lo largo de todos los codones y de todas las secuencias que son comparadas. Sin embargo, este modelo resulta limitado para detectar selección debido a que las secuencias codificantes suelen presentar regiones con tasas evolutivas variables. Por ello, se desarrollaron modelos capaces de estimar un ω para cada codon de la secuencia (Yang *et al.* 2000a). Estos modelos son por tanto capaces de detectar selección positiva en aminoácidos concretos. Los métodos que calculan diferentes valores de ω entre codones o entre especies (ver abajo) requieren, además del alineamiento de secuencias (parámetro *seqfile* en el archivo control de codeml), la introducción de una filogenia de las especies analizadas (parámetro *treefile* y con *runmode* = 0) como información de entrada.

Para inferir selección positiva, se compararon una serie de modelos que asumen un valor de ω entre 0 y 1 (modelos neutrales) con modelos que incluyen una clase adicional de $\omega > 1$ (modelos de selección). En primer lugar, se comparó el modelo neutral M1a, que asume dos clases de valores ($0 < \omega_0 < 1$; $\omega_1 = 1$) con el modelo de selección M2a, con tres clases de valores ($0 < \omega_0 < 1$; $\omega_1 = 1$; $\omega_2 > 1$). En segundo lugar, se comparó un modelo neutral que asume una distribución beta para los

valores de ω (M7) con un modelo similar pero que incluye una clase de $\omega > 1$ (M8). Por último, el modelo M8 se comparó con él mismo pero fijando $\omega = 1$ en la clase adicional. Este análisis reduce la probabilidad de obtener falsos positivos (Swanson *et al.* 2003).

En codeml, estos análisis se realizaron configurando los parámetros *model* = 0 y *nsites* = 1 (M1a), 2 (M2a), 7 (M7), 8 (M8). Para el modelo M8a, con *nsites* = 8, se fijó *fix_omega* = 1 y *omega* = 1.

Dado que los diferentes modelos dan su propio valor de probabilidad (likelihood), el análisis estadístico LRT (Likelihood Ratio Test) compara los valores de probabilidad entre los modelos neutrales y de selección mediante la siguiente fórmula:

$$\text{LRT} = 2x (\log\text{-likelihood}_{\text{selección}} - \log\text{-likelihood}_{\text{neutral}})$$

El valor P del estadístico resultante se estima bajo una distribución chi-cuadrado con el número de grados de libertad equivalente a la diferencia en el número de parámetros entre los dos modelos.

En los casos en los que el LRT fue significativo, dando una mayor probabilidad a los modelos de selección, se realizó un Análisis Bayesiano Empírico (BEB) para inferir los sitios bajo selección positiva, que son aquellos que presentan una probabilidad posterior > 0.95 de tener $\omega > 1$ (Yang *et al.* 2005).

3.5.2. *Estimación de tasas evolutivas específicas para cada linaje*

Además de asumir una única tasa evolutiva para todos los codones de una secuencia, el modelo de Goldman y Yang (1994) también asume que las tasas evolutivas no varían entre las diferentes especies que se analizan. No obstante, la evolución de una secuencia puede cambiar a lo largo del tiempo y por tanto puede presentar diferentes cambios entre las especies. Por tanto, se han desarrollado los modelos de rama libre, que estiman un valor de ω promedio para cada linaje en una filogenia (Yang y Nielsen 1998), pudiendo comparar la tasa evolutiva entre varias especies.

En codeml, el modelo de rama libre se aplica cambiando el parámetro *model* = 1 para estimar un ω diferente para cada especie. Si se selecciona *model* = 2, se puede calcular un valor de ω para un subconjunto de ramas marcadas en el código del árbol con el símbolo “#”, o bien “\$” si lo que se pretende marcar es un clado entero.

Los ω calculados en las ramas terminales de cada linaje solo representan los cambios producidos desde el ancestro más reciente. Sin embargo, los datos fenotípicos con los que se compararían los valores de ω reflejan la evolución del carácter a lo largo de toda la filogenia. Para corregir esta diferencia de escalas evolutivas entre datos genotípicos y fenotípicos, se recalcularon los valores de ω desde el ancestro más común. Esto se llevó a cabo sumando los valores de d_N y d_S desde la raíz del árbol hasta la rama terminal de cada especie y calculando el cociente entre ambos. De este modo, se tienen en cuenta todos los cambios acumulados por la región de estudio durante su

evolución, que es más adecuado para evaluar relaciones con datos fenotípicos (Montgomery *et al.* 2011). Además, la estimación de las tasas evolutivas desde el ancestro común presenta la ventaja de que las longitudes de las ramas se igualan, por lo que la filogenia no se encuentra sujeta a efectos temporales (2009). El cálculo de ω desde la raíz hasta cada rama terminal fue llevado a cabo usando la herramienta ETE2 (Huerta-Cepas *et al.* 2010) para Python.

3.5.3. *Estimación de tasas evolutivas específicas para cada linaje y entre codones*

Además de los métodos que estiman diferentes ω entre codones de una secuencia o entre especies de una filogenia, se han desarrollado otros modelos que permiten variar la tasa evolutiva tanto entre sitios como entre linajes. Estos modelos por tanto permiten la detección de selección positiva en especies determinadas (Yang y Nielsen 2002; Zhang *et al.* 2005).

El método está basado en dividir la filogenia de especies entre un *foreground* y un *background*, de tal modo que solo es posible detectar selección positiva en los linajes del *foreground*. Un LRT entonces se usa para comparar la probabilidad (*likelihood*) de un modelo nulo con la de un modelo alternativo. En el modelo alternativo, los codones son asignados a cuatro clases posibles:

Clase 0: a ella se asignan los sitios más conservados en todos los linajes con $0 < \omega_0 < 1$; Clase 1: contiene los sitios bajo evolución neutral ($\omega_1 = 1$) en todos los linajes.

Clase 2a: incluye los sitios bajo selección positiva $\omega_2 > 1$ en el *foreground* y con $0 < \omega_2 < 1$ en el *background*.

Clase 2b: incluye los sitios bajo selección positiva $\omega_2 > 1$ en el *foreground* y con $\omega_2 = 1$ en el *background*.

El modelo alternativo es entonces comparado con el modelo nulo, que es similar pero con ω_2 fijado a 1 en las clases 2a y 2b.

En codeml (PAML) este modelo es configurado asignando *model* = 2 y *nsites* = 2 para el modelo alternativo, mientras que para llevar a cabo el modelo nulo hay que seleccionar *fix_omega* = 1 y *omega* = 1. El *foreground* se selecciona insertando un símbolo “#” en el código del árbol filogenético junto a la rama que se quiere analizar. En el caso de evaluar si existe selección positiva en un subgrupo de linajes, se puede marcar el clado con el símbolo “\$”.

3.5.4. *Análisis de coevolución*

Para evaluar si dos proteínas muestran tasas evolutivas correlacionadas, se llevó a cabo una regresión entre los valores de ω estimados para cada gen. Los valores de ω se calcularon para cada rama de la filogenia usando el modelo de codeml descrito en el apartado 3.5.2. También se estimaron los valores de ω desde la raíz hasta cada rama terminal. Las tasas evolutivas de los diferentes genes fueron correlacionadas mediante un análisis de regresión lineal usando el programa estadístico R (<http://www.r-project.org/>) y los valores de los diferentes genes fueron utilizados como variables predictoras y dependientes en análisis por separado.

3.6. MEDIDA DE LOS NIVELES DE COMPETICIÓN ESPERMÁTICA Y DE LOS PARÁMETROS ESPERMÁTICOS

3.6.1. *Índice de competición espermática*

Una respuesta generalizada en las especies ante un incremento en los niveles de competición espermática es el incremento de la masa testicular con respecto a la masa corporal. Este fenómeno se ha observado en taxones muy diversos (Briskie y Robert 1992; Jennions y Passmore 1993; Gage 1994; Stockley *et al.* 1997; Gomendio *et al.* 1998). Además, en especies que producen varias crías por camada se ha detectado que el aumento del tamaño relativo de testículos se encuentra correlacionado con un alto índice de paternidad (Soulsbury 2010). Por lo tanto, la masa testicular en relación a la masa corporal se considera un índice fiable de los niveles de competición espermática. En los análisis de regresión, la masa testicular corregida por la masa corporal en un análisis de regresión múltiple (masa testicular residual) se utilizó como variable predictora para tratar de encontrar correlaciones entre la competición espermática y los parámetros evolutivos de los genes reproductivos. La masa relativa de los testículos como tal también se calculó utilizando la ecuación para roedores $Y = 0,031 X^{0,77}$ (Kenagy y Trombulack 1986), donde Y es la masa en gramos de ambos testículos y X es el peso corporal en gramos. Aunque este índice no se utilizó en los análisis estadísticos sí se usó para la representación gráfica de algunos resultados. Los valores de masa corporal, masa testicular y masa relativa de testículos para cada una de las especies utilizadas en este trabajo fueron aportadas por Laura Gómez Montoto y Maximiliano Tourmente, del Museo Nacional de Ciencias Naturales (CSIC). Los datos se muestran en la Tabla IV.

Tabla IV. Valores de masa corporal, masa testicular y masa relativa de testículo de las especies utilizadas en este trabajo.

Especie	Masa corporal (g)	Masa testicular (g)	Masa relativa de testículo
<i>Apodemus sylvaticus</i>	30,46	1,08	2,52
<i>Arvicola sapidus</i>	217,67	2,24	1,14
<i>Arvicola terrestris</i>	91,56	0,41	0,41
<i>Chionomys nivalis</i>	50,33	1,03	1,62
<i>Clethrionomys glareolus</i>	29,27	0,63	1,50
<i>Cricetulus griseus</i>	33,72	1,78	3,83
<i>Lemniscomys barbarus</i>	44,62	0,67	1,15
<i>Mastomys natalensis</i>	81,17	0,92	1,01
<i>Mesocricetus auratus</i>	124,99	3,50	2,75
<i>Microtus arvalis</i>	44,26	0,44	0,77
<i>Microtus cabreræ</i>	45,32	0,15	0,26
<i>Pitimys duodecimcostatus</i>	29,55	0,08	0,19
<i>Pitimys lusitanicus</i>	17,73	0,09	0,33
<i>Mus m. bactrianus</i>	18,06	0,18	0,61
<i>Mus caroli</i>	18,07	0,14	0,50
<i>Mus m. castaneus</i>	18,82	0,08	0,26
<i>Mus cookii</i>	23,66	0,31	0,86
<i>Mus domesticus</i>	22,05	0,11	0,32
<i>Mus famulus</i>	27,40	0,05	0,13
<i>Mus macedonicus</i>	20,10	0,30	0,95
<i>Mus minutoides</i>	5,54	0,10	0,90
<i>Mus m. musculus</i>	21,85	0,14	0,43
<i>Mus pahari</i>	33,15	0,13	0,28
<i>Mus spicilegus</i>	18,10	0,43	1,49
<i>Mus spretus</i>	18,17	0,30	1,04
<i>Phodopus campbelli</i>	48,52	1,93	3,16
<i>Phodopus roborovskii</i>	25,72	1,06	2,82
<i>Phodopus sungorus</i>	46,70	0,99	1,67

3.6.2. *Parámetros de velocidad espermática*

Los parámetros de velocidad espermática se midieron usando el sistema de análisis computarizado Sperm Class Analyzer v.4.0 (Microptic, Barcelona, Spain). Un total de 5 µl de suspensión de espermatozoides en medio mT-H se colocaron en un portaobjetos con una cámara de 20-µm de profundidad (Standard Count-2 Chamber Slide 20-micron, Leja, Nieuw-Vennep, Netherlands) pre-calentada a 37°C, y se examinaron en un microscopio de contraste de fases con objetivo 4x. Las medidas se realizaron durante los 5 minutos posteriores a la colección de las muestras.

Utilizando una cámara de video (Basler A312fc, Vision Technologies), se grabaron hasta ocho videos de 4 segundos para cada muestra de espermatozoides. La concentración de espermatozoides

se ajustó previamente a $4-6 \times 10^6$ espermatozoides/ml para satisfacer los requerimientos de los análisis. Los videos se analizaron obteniendo un mínimo de 150 tomas para cada muestra. Para los análisis se utilizaron 5 individuos de cada especie.

Se cuantificaron un total de siete parámetros de velocidad espermática: velocidad curvilínea (VCL) (en $\mu\text{m/s}$), velocidad de la trayectoria media (VAP) (en $\mu\text{m/s}$), velocidad rectilínea (VSL) (en $\mu\text{m/s}$), amplitud del desplazamiento lateral de la cabeza (ALH) (en $\mu\text{m/s}$), frecuencia de batido (BCF) (en Hz), linealidad ($\text{LIN} = (\text{VSL}/\text{VCL}) \times 100$) (en %), rectitud ($\text{STR} = (\text{VSL}/\text{VAP}) \times 100$) (en %) y (Fig.). Con el fin de reducir componentes de velocidad ya correlacionados al mínimo número de factores representando toda la información, se llevó a cabo un análisis de componentes principales (PCA).

Los parámetros de velocidad espermática de las diferentes especies de roedores fueron medidos por Laura Gómez Montoto y Maximiliano Tourmente, del Museo Nacional de Ciencias Naturales (CSIC).

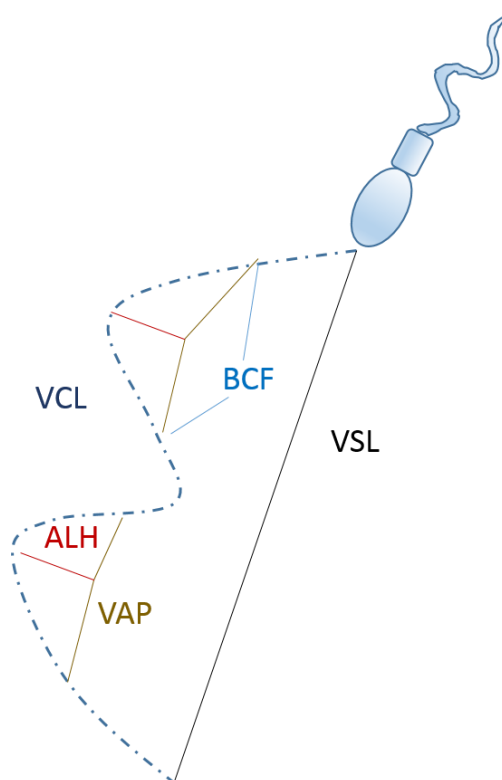


Figura 7. Descriptores de velocidad espermática medidos por el sistema SCA. VCL: velocidad curvilínea; VSL: velocidad rectilínea; VAP: velocidad de la trayectoria media; ALH: amplitud del desplazamiento lateral de la cabeza; BCF: frecuencia de batido.

3.7. ANÁLISIS DE CORRELACIÓN CORREGIDOS POR EFECTOS FILOGENÉTICOS

Cuando se llevan a cabo análisis de regresión en los que se comparan diferentes rasgos entre especies, se debe tener en cuenta que los datos no son independientes entre sí, ya que mantienen una relación filogenética. Por ello, es importante poder controlar los efectos de la filogenia sobre los análisis estadísticos. De este modo, se puede garantizar que las diferencias observadas entre los linajes son producto de la variación biológica y que no se encuentran influidas por su distancia filogenética. Por tanto, los análisis estadísticos se llevaron a cabo utilizando la metodología de regresión de mínimos cuadrados con corrección filogenética (PGLS). Los análisis de PGLS se realizaron con la librería Caper (Orme *et al.* 2012) implementada en el programa estadístico R (<http://www.r-project.org/>). Durante el análisis se estima el parámetro filogenético lambda (λ), que indica el nivel de ajuste de nuestros datos a un modelo de evolución Browniano. Un valor de λ próximo a cero indica que las variables han evolucionado independientemente de la filogenia. Si por el contrario el valor de λ es próximo a 1 nos indica que las variables presentan una fuerte asociación filogenética entre ellas.

3.8. ANÁLISIS GENÓMICOS EN CEPAS DE RATÓN

3.8.1. Selección de proteínas

Se recopiló una base de datos de 1.350 proteínas presentes en el espermatozoide de ratón utilizando las listas de proteomas publicados previamente (Cao *et al.* 2006; Stein *et al.* 2006; Baker *et al.* 2008a; Dorus *et al.* 2010). De entre todas estas, para los análisis evolutivos se seleccionaron 165 proteínas que presentan una función confirmada o fuertemente apoyada en alguno de los procesos que tienen lugar en el espermatozoide. La información sobre las funciones de las proteínas fue recogida de diferentes bases de datos, como UniProt (<http://www.uniprot.org>), Mouse Genome Informatics (<http://www.informaticsjax.org>), y AmiGO (<http://www.geneontology.org>), así como de una revisión exhaustiva de la literatura. Las proteínas se clasificaron en seis categorías en función del proceso en el que participan dentro del espermatozoide: espermatogénesis, metabolismo, motilidad, capacitación, reacción acrosómica e interacción espermatozoide-óvulo. Cada proteína se asignó a un solo grupo para garantizar la independencia estadística entre categorías.

3.8.2. Obtención de secuencias de los genomas de ratón

Las secuencias de los genes que codifican para las proteínas seleccionadas se extrajeron de los genomas de un total de 17 cepas de ratón (Fig. 8). Estas incluyen 13 cepas clásicas de laboratorio (129SvEv, 129P2/OlaHsd, 129S1/SvImJ, C3H/HeJ, CBA/J, A/J, AKR/J, DBA/2J, LP/J, BALB/cJ, NZO/HILtJ y NOD/ShiLtJ), y cuatro cepas derivadas de especies silvestres (PWK/PhJ, WSB/EiJ

CAST/EiJ and SPRET/EiJ), las cuales representan las especies de *M. musculus musculus*, *M. m. domesticus*, *M. m. castaneus* y *M. spretus*, respectivamente.

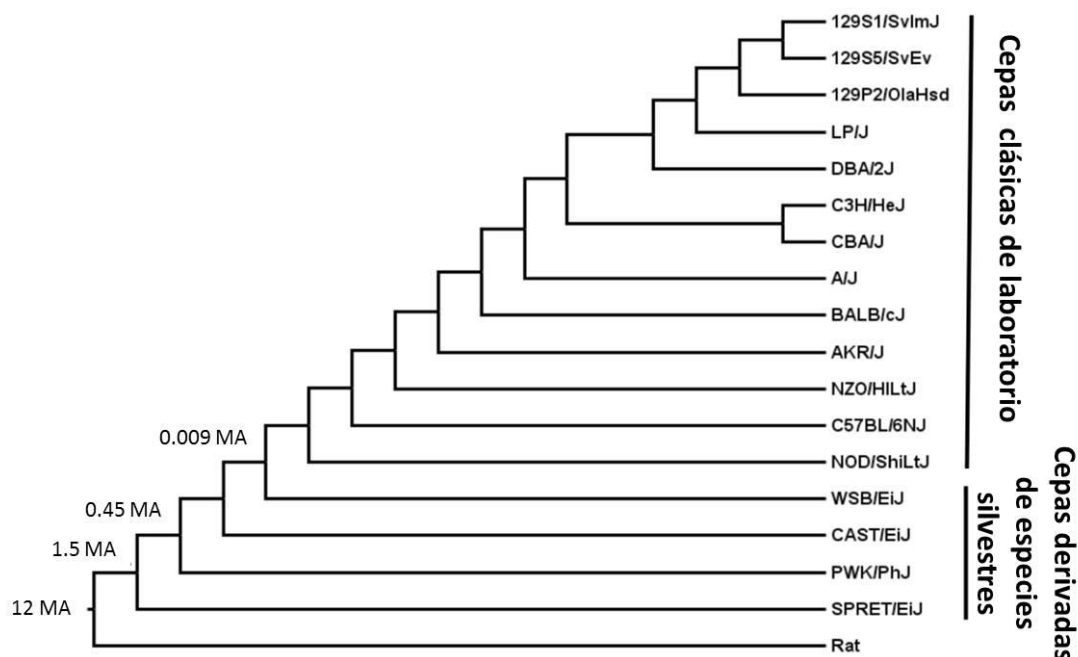


Figura 8. Cladograma de las cepas de ratón. El tiempo de divergencia en millones de años (MA) se muestra para las separaciones entre cepas derivadas de especies silvestres así como para la separación entre la cepa derivada de *Mus m. domesticus* (WSB/EiJ) y las cepas de laboratorio. Rata se utilizó como grupo externo. El cladograma se construyó basado en información de la literatura (Beck *et al.*2000; Suzuki *et al.*2004; Goios *et al.*2007; Keane *et al.*2011).

Los genomas consenso fueron descargados en formato FASTA del servidor FTP del Mouse Genomes Project del Instituto Sanger (ftp://ftp-mouse.sanger.ac.uk/current_consensus/) (Keane *et al.* 2011). Las secuencias génicas fueron extraídas de los diferentes genomas en tándem usando un script escrito en lenguaje Perl. Las regiones intónicas y UTRs fueron localizadas y descartadas basándonos en la anotación de las secuencias de Ensembl (<http://www.ensembl.org/>). Los exones resultantes fueron ensamblados para reconstruir la secuencia codificante usando el programa BioEdit. Las secuencias codificantes se alinearon utilizando el programa ClustalW.

3.8.3. Análisis estadísticos

Sobre los parámetros evolutivos estimados de los análisis de codeml (d_N , d_S , ω) se aplicó el test de Shapiro-Wilks para comprobar que tenían distribuciones normales. Debido a que ninguna de las variables cumplía el principio de normalidad, las comparaciones estadísticas de las distribuciones de d_N , d_S y ω entre las diferentes categorías de procesos reproductivos se llevaron a cabo mediante análisis de Kruskal-Wallis no paramétricos. El programa de estadística utilizado para estos análisis fue InfoStat (<http://www.infostat.com.ar>).

Dado que estos análisis estadísticos realizan múltiples comparaciones de manera simultánea, esto puede conllevar a rechazar la hipótesis nula de manera errónea (error de tipo I). Por tanto, para corregir el error por la multiplicidad de análisis se aplicó la función *KruskalMC* (Siegel 1988) implementada en el paquete *Pgirmess* del programa de estadística R (<http://www.r-project.org/>).

Las proporciones de los genes que se encuentran bajo selección positiva en cada categoría se compararon a las proporciones esperadas tras aplicar una chi-cuadrado con 5 grados de libertad y un nivel de significancia de $p = 0,01$. Los valores esperados se calcularon como la proporción de proteínas dentro de cada categoría respecto al número total de proteínas.

3.9. ANÁLISIS DE PROTEÓMICA DEL ESPERMATOZOIDE

3.9.1. Preparación de muestras

Se sacrificó a machos de ratón por dislocación cervical y los espermatozoides se tomaron del epidídimo tras dejarlos difundir en 1-3 ml de medio Dubecco salino tamponado con fosfato (DPBS) (Gibco, Madrid, Spain).

Las muestras de espermatozoides provenían de un pool de tres individuos por especie para tener en cuenta la variabilidad biológica y asegurar que se alcanzaba la concentración de proteínas necesaria (Karp *et al.* 2009). Los espermatozoides fueron diluidos en 900 μ l de DPBS y centrifugados a 6.000 rpm durante 1 min. Después de tres lavados, se descartó el sobrenadante y el pellet se guardó a -80°C hasta su uso.

3.9.2. Cuantificación de las muestras

Los pellets se solubilizaron en 15 μ l de buffer de conservación conteniendo agente reductor 1 x Nupage® (Life Technologies, Paisley, Reino Unido). La cantidad de proteína se determinó usando el kit comercial EZQ Protein Quantitation (Life Technologies) siguiendo instrucciones del fabricante. La fluorescencia de proteína se midió con un scanner de laser FLA-2000 (Fuji Film) con 473 nm de excitación y filtros de emisión LPR. EL programa AIDA v 3.23 fue utilizado para analizar la fluorescencia y producir las curvas de cuantificación y los valores de estimación.

3.9.3. Electroforesis diferencial en geles de 2 dimensiones (2D-DIGE)

2D-DIGE es una técnica de proteómica cuantitativa que consiste en comparar en un mismo gel 2D dos muestras de proteínas marcadas con sondas fluorescentes distintas. Este método por tanto permite detectar diferencias en abundancia de proteínas pertenecientes a diferentes muestras, o bien entre muestras con diferentes tratamientos. En nuestro caso, el experimento que se llevó a cabo fue comparar muestras de proteínas del espermatozoide de tres especies de roedores múridos con diferentes niveles de competición espermática: *M. musculus*, *Mus spretus* y *Mus spicilegus*.

Para llevar a cabo los análisis comparativos en geles 2D, se siguió el protocolo basado en el uso de dos sondas de Karp *et al.* (2007). Las muestras fueron marcadas con sondas fluorescentes de Cianina (CyDye, GE Healthcare, Cambridge, Reino Unido) siguiendo el procedimiento de Karp y Lilly (2005). Las muestras marcadas con sondas CyDye fueron cargadas en tiras IPG de 13 cm pH 3-10) durante 10 h a 20°C y 20 V usando el aparato IPGphor II (GE Healthcare). El isoelectroenfoque (IEF) fue llevado a cabo aplicando 40000 Va 20°C y 10 mA. Antes de la SDS-PAGE, las tiras se equilibraron durante 15 min en Tris pH 6.8 100 mM, glicerol 30%, urea 8 M, SDS1%, azul de bromofenol 0,2 mg/ml. Las tiras fueron cargadas en un gel de acrilamida al 12%, pH 8,8, de 13 cm de ancho y 1mm de grosor. Las tiras se cubrieron previamente con agarosa al 1% en tampón SDS (Tris 25 mM pH 8.3, glicina 192 mM, y SDS al 0.1%) conteniendo 5 mg de azul de bromofenol. Los geles corrieron durante 15 min a 20 mV y entonces se dejó a 40mV hasta que el frente de bromofenol alcanzó la parte inferior del gel. Se prepararon dos réplicas de geles para cada especie.

Los *spots* de los geles se visualizaron usando el programa Typhoon™ 9410 (GE Healthcare) y se analizaron con DeCyder BVA V5.0 (GE Healthcare), siguiendo las recomendaciones de fabricante. Los datos de abundancia para cada *spot* se normalizaron usando un método radiométrico y se transformaron a escala logarítmica para estabilizar la varianza. Los datos de abundancia entre geles se compararon estadísticamente mediante la pruebas de la t de Student y de análisis de la varianza (ANOVA). Las comparaciones con un ratio promedio absoluto > 1.5 y con una $p < 0,05$ fueron seleccionados para los análisis de espectrometría de masas. La validez de estas diferencias en abundancia fue confirmada tras comprobar visualmente la intensidad de los *spots* en los geles.

Los *spots* candidatos fueron tomados manualmente del gel y digeridos con 12.5 ng/ml de tripsina de secuencia (Roche Applied Science) en un buffer de Tris-HCl 20 mM, pH 8,0, y CaCl₂ 5 mM a 37°C. Los péptidos trípticos se extrajeron del gel con una solución de acetonitrilo 60% (Romil) y ácido fórmico ARISTAR 4%. 0,5 µl de péptidos se mezclaron con 0.5 µl de una matriz de ácido α -cyano-4-hidroxicinámico y 0,5 µl de esta mezcla se motearon sobre una placa para ser analizada por el espectrómetro de masas MALDI-TOF-TOF 4800 (Applied Biosystems). El aparato se calibró con productos bovinos de autólisis y una matriz de iones de calcio. Los datos sobre masas de péptidos y de fragmentación fueron buscados en la base de datos de RefSeq para la especie *M. musculus* usando MASCOT, validando todas las entradas con una tolerancia de péptidos de ± 25 ppm, tolerancia de espectrometría de masas de $\pm 0,8$ Da, permitiendo rupturas de uno, y usando NEM como modificación variable.

3.9.4. Cromatografía líquida y espectrometría de masas cuantitativa

Además del método 2D-DIGE, en nuestro estudio utilizamos otra metodología de proteómica cuantitativa que no requiere del marcaje de las proteínas. Este tipo de técnicas consisten en el uso de la cromatografía líquida seguida de espectrometría de masas (LC-MS/MS) para la estimación de la

abundancia de las proteínas, la cual se basa en la cuantificación del número de péptidos detectados para cada proteína. Esta técnica presenta ciertas ventajas frente al uso de geles 2D ya que tiene un rango dinámico más amplio, evita los problemas relacionados con la saturación del marcaje, y ofrece una mayor resolución.

Un total de 50 µg de muestra de espermatozoides se separaron en geles Nupage® Bis-Tris al 4-12 % en un sistema XCell SureLock™ Mini-Cell PowerEase 200 (Life-Technologies), siguiendo las condiciones indicadas por el fabricante. Los geles entonces se fijaron en una solución de metanol 45% y ácido acético 1% durante 1 h e inmediatamente se marcaron con azul de Coomassie (Coomassie 0,1% v/v, metanol 34% v/v, sulfato de amonio 17% v/v, y ácido acético al 0,5% v/v) durante 12 h hasta que las bandas de proteína fueron visibles. Los geles se transfirieron a una rebanadora donde cada línea fue cortada en 16 pedazos. Los pedazos fueron sujetos a digestión proteolítica con el sistema Automated Preparation Station (Perkin Elmer, Cambridge, Reino Unido). Las proteínas fueron reducidas y alcalinizadas y entonces se trataron con ditiotreitól y yodoacetamida, seguidos de digestión por tripsina 500 ng/µl (Promega, Cambridge, Reino Unido) durante 16 h a 37 °C. Tras la digestión, 5 µl de cada muestra se inyectaron en una columna de cromatografía líquida de alta presión (HPLC) C-18 PepMap con un diámetro interno de 15cm x 75mm (Thermo, San Jose, CA) acoplado a un sistema Eksigent LC.

El eluido de la columna se introdujo en un espectrómetro de masas QExactive™ Hybrid Quadrupole-Orbitrap (Thermo) para el análisis y la cuantificación de los péptidos. Los datos crudos de la MS/MS se analizaron por X! Tandem (Craig y Beavis 2004) utilizando la base de datos del proteoma de ratón GPM. Los resultados de tres réplicas por especie se validaron por PeptideProphet™ y la asignación de proteínas se hizo con ProteinProphet™, ambas herramientas del Trans-Proteomic Pipeline (TPP) (Keller *et al.* 2005). Las asignaciones de proteínas con una tasa de falso descubrimiento mayores del 1% e identificados por un único péptido fueron descartadas.

Los análisis de 2D-DIGE y de LC-MS/MS fueron llevados a cabo en el Centro de Proteómica de la Universidad de Cambridge (Reino Unido).

3.9.5. *Cuantificación de proteínas con APEX*

Un problema que suelen presentar las técnicas de proteómica basadas en la cuantificación del número de espectros es que cada péptido presenta unas propiedades fisicoquímicas determinadas, que provoca que la probabilidad de detección de los péptidos por espectrometría de masas sea diferente. Esto implica que en muchos casos las abundancias estimadas mediante cuantificación de espectros no se corresponden con las abundancias reales de las proteínas en la muestra. Para evitar este inconveniente se ha desarrollado APEX (Absolute Protein Expression) (Lu *et al.* 2007; Braisted *et al.* 2008), que es un método computacional de cuantificación que corrige la abundancia observada

(número de espectros) con valores de abundancia esperada para obtener una estimación de la abundancia real de la proteína en la muestra.

Para calcular la abundancia esperada de las proteínas, APEX predice la probabilidad de detección de los péptidos en función de sus propiedades fisicoquímicas usando un algoritmo de clasificación. Este algoritmo tiene que ser previamente entrenado por un conjunto de proteínas de propiedades conocidas. En nuestro caso, se seleccionaron las 50 proteínas con el mayor número de espectros detectados y con la mayor probabilidad de detección dentro del espermatozoide de ratón, basado en un previo estudio de proteómica (Dorus *et al.* 2010). Las 35 propiedades fisicoquímicas disponibles en APEX fueron utilizadas para predecir la detección o no detección de los péptidos.

Las probabilidades de cada proteína (O_i) se computaron usando el algoritmo de clasificación Random Forest. Para los análisis de APEX, los datos de MS de las tres réplicas se combinaron en un solo set de datos, que se usó para calcular un único valor de abundancia APEX por especie. Se estableció una tasa de falsos positivos < 1% y no se aplicó un factor de normalización a los valores de APEX. Los datos de abundancia de APEX se compararon entre especies mediante el cálculo de un Z-score (Lu *et al.* 2007).

El riesgo a obtener errores de tipo I a causa de la multiplicidad de análisis fue prevenido con la aplicación de una corrección de Bonferroni.

3.9.6. Clasificación funcional de proteínas del espermatozoide

Los componentes del proteoma se clasificaron en función de su función biológica y molecular, así como de su localización celular, utilizando el sistema de clasificación PANTHER (Mi y Thomas 2009) (<http://www.pantherdb.org/>). La identificación de clases funciones enriquecidas en proteínas identificadas se llevó a cabo con la herramienta GOrilla (Eden *et al.* 2009) (<http://cbl-gorilla.cs.technion.ac.il/>) usando un nivel de significancia de $p = 10^{-5}$. Para estos análisis se utilizó como background el proteoma completo de ratón (38561 proteínas), al cual se tuvo acceso mediante la herramienta BioMart de Ensembl (<http://www.ensembl.org/biomart/martview/>).

La información relativa a los fenotipos reproductivos asociados con las proteínas analizadas se obtuvo de la base de datos de fenotipos mutantes del Mouse Genome Informatics (<http://www.informatics.jax.org/>). La anotación funcional de aquellas proteínas presentando diferencias significativas en abundancia se realizó con la herramienta DAVID (<http://david.abcc.ncifcrf.gov/>).

4. RESULTADOS

BLOQUE I

Las proteínas del espermatozoide de ratón implicadas en motilidad e interacción de gametos presentan una elevada tasa evolutiva.

Se ha planteado que los cambios adaptativos que se observan en diversos rasgos del espermatozoide son causados por la rápida divergencia de las proteínas reproductivas. Sin embargo, todavía no se comprenden bien las fuerzas selectivas que dirigen la evolución de las proteínas reproductivas. En cualquier caso, una variedad de fuerzas selectivas pueden afectar en diferente grado a las diversas etapas que tienen lugar desde la generación de los gametos hasta el momento de la fecundación. Por lo tanto, es posible que exista una adaptación compartimentada de las proteínas del espermatozoide en términos de su función molecular, su localización en la célula, o el proceso biológico en el que participan.

En este trabajo se analizó la evolución de un amplio grupo de proteínas implicadas en diferentes procesos que tienen lugar en el espermatozoide de ratón. Para ello, se utilizaron 17 cepas de ratón cuyos genomas han sido secuenciados. Cuatro de estas cepas derivan de las especies silvestres *Mus musculus musculus*, *Mus m. domesticus*, *Mus m. castaneus* y *Mus spretus*. Los análisis evolutivos revelaron una mayor proporción de proteínas relacionadas con la motilidad espermática y la interacción espermatozoide-óvulo exhibiendo una señal de selección positiva respecto a las otras categorías. Además, las proteínas implicadas en la interacción espermatozoide-óvulo muestran una evolución más acelerada en comparación con aquellas implicadas en otros eventos. Estos análisis además permitieron identificar un amplio grupo de proteínas candidatas para futuros estudios de asociación genotipo-fenotipo y para análisis comparativos entre especies sujetas a diferentes presiones de selección sexual.

4.1. Distribución de proteínas

Tras una revisión exhaustiva de diversas bases de datos de proteínas y de la bibliografía sobre proteómica del espermatozoide, se recopilaron un total de 1.350 proteínas que se han detectado hasta el momento en el espermatozoide de ratón. Dentro de éstas, para 165 proteínas se ha identificado una función en alguno de los procesos del espermatozoide que conducen a la fecundación (Tabla suplementaria 1). Un total de 33 (20%) proteínas se asociaron con espermatogénesis, 23 (13,93%) con metabolismo del espermatozoide, 34 (20,6%) con motilidad, 21 (12,7%) con capacitación, 20 (12,1%) con reacción acrosómica, y 34 (20,6%) con interacción espermatozoide-óvulo. La distribución de estas proteínas entre las diferentes categorías funcionales no mostró un sesgo respecto a la distribución funcional que presenta el proteoma del espermatozoide ($\chi^2_{16 \text{ g.l.}} = 0,898$, $p = 0,997$).

4.2. Tasas evolutivas de las proteínas del espermatozoide

Para las 165 proteínas seleccionadas, se extrajeron sus secuencias codificantes del genoma de 17 cepas de ratón (*ver* Fig. 10). Tras su alineamiento, la tasa evolutiva de las secuencias codificantes se calculó como el cociente entre cambios no sinónimos y cambios sinónimos ($d_N/d_S = \omega$). La estimación de los parámetros evolutivos entre las 165 proteínas resultó en un d_N promedio de 0,064 (DE = 0,089), un d_S promedio de 0,263 (DE = 0,29) y un ω promedio de 0,259 (DE = 0,255). La mayoría de las proteínas mostraron valores de ω entre 0 y 1 y solo se estimaron valores mayores que 1 para 6 proteínas (Tmem190, Hils1, Smcp, Slx11, Fscb y Crisp1) (Tabla suplementaria 2).

Las distribuciones de los parámetros evolutivos entonces se compararon entre las diferentes categorías de procesos reproductivos. Se observó un mayor valor promedio de ω para las proteínas implicadas en la interacción espermatozoide-óvulo ($\omega = 0,413$, DE = 0,304) en comparación con las otras categorías. Las proteínas relacionadas con el metabolismo del espermatozoide ($\omega = 0,156$, DE = 0,14), capacitación ($\omega = 0,168$, DE = 0,181) y la reacción acrosómica ($\omega = 0,178$, DE = 0,186) mostraron los valores promedio de ω más bajos. Las proteínas con un papel en espermatogénesis ($\omega = 0,262$, DE = 0,287) y motilidad espermática ($\omega = 0,266$, DE = 0,247) exhibieron valores promedio de ω intermedios (Fig. 9A). Los análisis estadísticos revelaron diferencias significativas entre la distribución de valores de ω de las diferentes categorías funcionales ($H = 23,18$, $p = 0,0003$). Las comparaciones entre pares de categorías revelaron diferencias significativas en los valores de ω de las proteínas implicadas en interacción espermatozoide-óvulo respecto al resto de categorías, con excepción de la motilidad espermática (Tabla V). Esta categoría no mostró diferencias significativas con ninguna otra.

Las proteínas de motilidad espermática presentaron el mayor valor promedio de cambios no sinónimos (d_N) ($d_N = 0,11$, DE = 0,143), seguido por las proteínas de interacción espermatozoide-óvulo ($d_N = 0,094$, DE = 0,089). Las estimas más bajas de d_N correspondieron a las categorías conservadas de metabolismo del espermatozoide ($d_N = 0,033$, DE = 0,022) y capacitación ($d_N = 0,031$, DE = 0,038) (Fig. 9B). Los análisis comparativos revelaron diferencias significativas en las distribuciones de d_N entre las diferentes categorías funcionales ($H = 28,29$, $p = 0,0002$). Las proteínas implicadas en motilidad espermática mostraron un valor promedio de d_N significativamente mayor que el resto de categorías excepto las proteínas asociadas con interacción espermatozoide-óvulo (Tabla VI).

También se detectaron diferencias significativas para las distribuciones de cambios sinónimos (d_S) entre las diferentes categorías de procesos reproductivos ($H = 17,96$, $p = 0,003$). Las proteínas implicadas en motilidad espermática mostraron el mayor valor promedio de d_S ($d_S = 0,418$, DE = 0,561), aunque los datos se presentaron muy dispersos, dando lugar a una amplia desviación típica (Fig. 9B). La motilidad espermática difirió significativamente de aquellas categorías con los valores

más bajos de d_s como espermatogénesis ($d_s = 0,182$, $DE = 0,076$) y capacitación ($d_s = 0,191$, $DE = 0,128$) (Tabla VI).

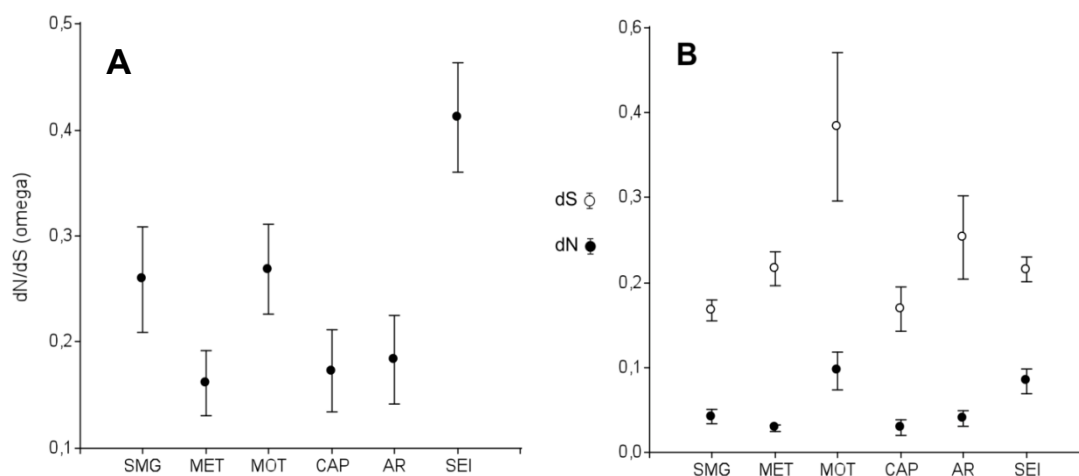


Figura 9. Tasas evolutivas de las proteínas de espermatozoide de ratón. Las comparaciones entre las diferentes categorías de procesos reproductivos son mostradas. **A.** comparación de los valores promedio de ω (dN/dS) y **B.** valor promedio de cambios no sinónimos (dN , círculos negros) y sinónimos (dS , círculos blancos). Abreviaturas de las categorías: espermatogénesis (SMG), metabolismo (MET), motilidad (MOT), capacitación (CAP), reacción acrosómica (AR) e interacción espermatozoide-óvulo (SEI). Los errores típicos son mostrados.

Diversos estudios previos han sugerido que las proteínas con un $\omega > 0,5$ probablemente experimentan evolución adaptativa (ej: *ver Swanson et al. 2004*). Por ello, se seleccionaron aquellas proteínas mostrando un $\omega > 0,5$ (Tabla suplementaria 2) y se observó su distribución entre las diferentes categorías. Un total de 25 proteínas con $\omega > 0,5$ fueron seleccionadas, y se detectaron diferencias significativas entre la proporción de estas proteínas dentro de cada categoría ($\chi^2_{5 \text{ g.l.}} = 22,2$, $p = 0,00047$). La categoría con mayor proporción de proteínas con $\omega > 0,5$ fue la de interacción espermatozoide-óvulo (9 de 25; 36%), con una desviación significativa respecto a su proporción esperada (Tabla VII).

Tabla V. Comparación de tasas evolutivas (ω) entre los diferentes procesos reproductivos.

Proceso reproductivo	Espermatogénesis	Metabolismo	Motilidad	Capacitación	Reacción acrosómica	Interacción espermatozoide-óvulo
Espermatogénesis	-	0,27	0,64	0,23	0,38	0,033 (*)
Metabolismo	-	-	0,12	0,90	0,86	0,0016 (**)
Motilidad	-	-	-	0,10	0,20	0,12
Capacitación	-	-	-	-	0,77	0,0014 (**)
Reacción acrosómica	-	-	-	-	-	0,0061 (**)
Interacción espermatozoide-óvulo	-	-	-	-	-	-

Los valores de p para cada comparación son mostrados. (*) y (**) indican significancia estadística con un nivel de significancia del 95% y un 99% respectivamente.

Tabla VI. Comparaciones de cambios no sinónimos (d_N) y sinónimos (d_S) entre los diferentes procesos reproductivos.

Proceso reproductivo	Espermatogénesis	Metabolismo	Motilidad	Capacitación	Reacción acrosómica	Interacción espermatozoide-óvulo
Espermatogénesis	-	0,88	0,06	0,30	0,86	0,0021 (**)
Metabolismo	0,05	-	0,06	0,64	0,54	0,0005 (**)
Motilidad	0,002 (**)	0,15	-	0,028 (*)	0,08	0,13
Capacitación	0,75	0,25	0,025 (*)	-	0,96	0,0003 (**)
Reacción acrosómica	0,05	0,93	0,18	0,21	-	0,024 (*)
Interacción espermatozoide-óvulo	0,09	0,92	0,14	0,09	0,94	-

Sobre la diagonal se representan los valores p resultantes de la comparación de d_N , y debajo de la diagonal los de d_S . (*) y (**) indican significancia estadística con un nivel de confianza del 95% y un 99% respectivamente.

Tabla VII. Representación de las proteínas con evolución rápida ($\omega > 0,5$) en los diferentes procesos reproductivos.

Proceso reproductivo	Número de proteínas con $\omega > 0,5$	Proporción observada	Proporción esperada	Observada - esperada
Espermatogénesis	6	24,0	20,0	4,0
Metabolismo	3	8,0	13,94	-5,94
Motilidad	4	20,0	20,61	-0,61
Capacitación	1	4,0	12,73	-8,73
Reacción acrosómica	2	8,0	12,12	-4,12
Interacción espermatozoide-óvulo	9	36,0	20,61	15,39

Las proporciones esperadas se calcularon como el porcentaje de genes en cada categoría respecto al número total de proteínas.

4.3. Selección positiva en las proteínas del espermatozoide

Aunque los valores globales de ω son indicadores importantes de la presión selectiva a nivel de proteína, muchas proteínas presentan una elevada proporción de aminoácidos que permanecen altamente conservados, y los episodios de selección suelen afectar a unos pocos residuos (Yang *et al.* 2000a).

Por lo tanto, se llevaron a cabo análisis evolutivos que permiten la detección de aminoácidos dentro de la proteína que se encuentran bajo selección positiva (Yang *et al.* 2000a). Se identificaron un total de 48 (29,1% del total) proteínas mostrando una señal de selección positiva (Tabla suplementaria 3). Se observaron diferentes proporciones de proteínas experimentando selección positiva entre las categorías. Un total de 10 proteínas bajo selección positiva pertenecían a la categoría de espermatogénesis (30,3%), 6 a metabolismo del espermatozoide (26,1%), 14 a motilidad espermática (41,2%), 4 a capacitación (19,0%), una a reacción acrosómica (20,5%) y 13 a interacción espermatozoide-óvulo (41,2%) (Fig. 10).

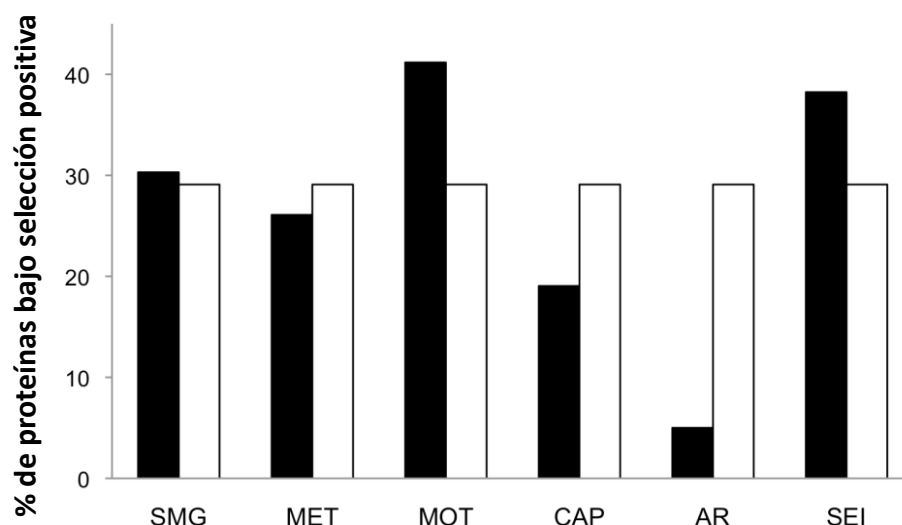


Figura 10. Selección positiva en proteínas del espermatozoide. La proporción de proteínas bajo selección positiva en cada categoría de proceso reproductivo es representada. Abreviaturas de las categorías: espermatogénesis (SMG), metabolismo (MET), motilidad (MOT), capacitación (CAP), reacción acrosómica (AR) e interacción espermatozoide-óvulo (SEI).

La proporción de proteínas sujetas a selección positiva difirió significativamente entre las categorías ($\chi^2_{5 \text{ g.l.}} = 31.7$, $p = 6,9 \times 10^{-6}$). Las categorías de motilidad espermática e interacción espermatozoide-óvulo superaron al porcentaje esperado de proteínas bajo selección positiva (29,1%), estimado como el porcentaje de proteínas exhibiendo selección positiva respecto al total.

Con el fin de identificar si la selección positiva actúa sobre dominios de proteínas importantes para la función del espermatozoide, se estudiaron tres ejemplos de proteínas implicadas en interacción espermatozoide-óvulo (Zan, Adam32 y Crisp1) con evidencia de selección positiva (Fig.11). En la zonadhesina (Zan), se detectó una señal intensa de selección positiva en el dominio mucin-like y un elevado número de sitios bajo selección a lo largo de las repeticiones parciales D3, las cuales se han generado por duplicaciones en tándem. En el caso de Adam32, 4 de los 6 aminoácidos bajo selección positiva se encuentran en los dominios de adhesión desintegrina/rico en cisteína, la región que se presume implicada en la interacción con las integrinas del óvulo. En Crisp1, los cinco aminoácidos bajo selección positiva se distribuyen entre el dominio CAP y la región C-terminal del dominio rico en cisteína, mientras que los 16 residuos de cisteína característicos de la familia de proteínas CRISP se encuentran conservados.

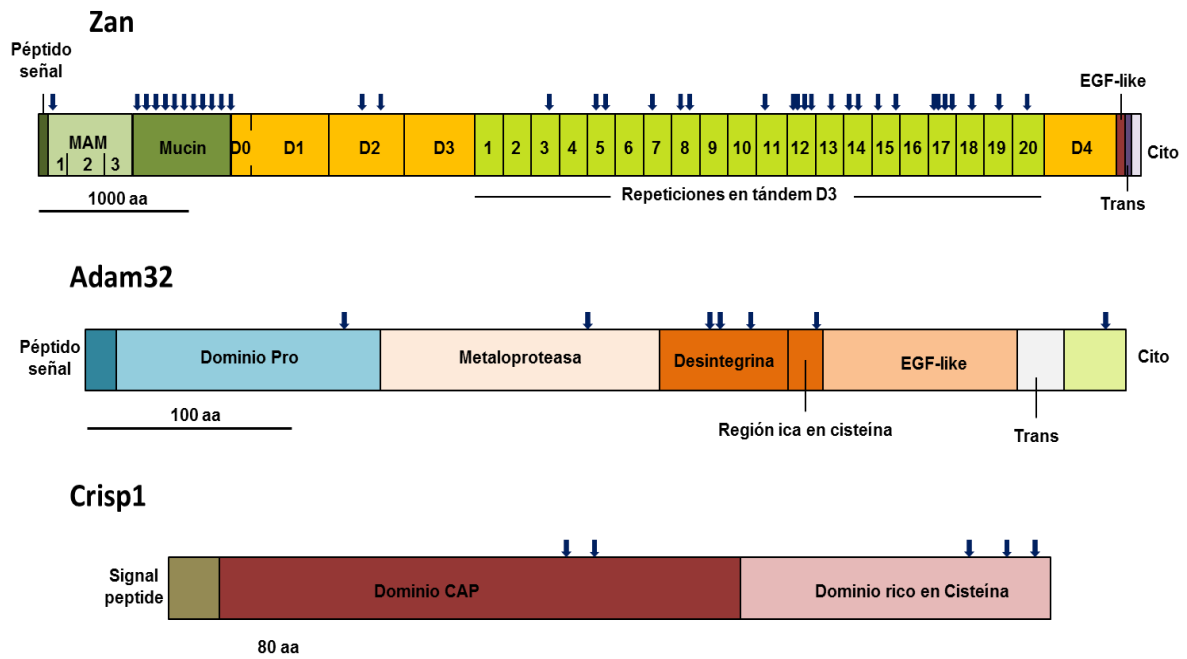


Figura 11. Representación esquemática de tres proteínas sujetas a evolución adaptativa. Los aminoácidos identificados bajo selección positiva con una probabilidad posterior mayor a 0,95 se indican con flechas. La escala relativa en aminoácidos se muestra para cada proteína. La organización de los dominios se basa en la información de UniProt.

BLOQUE II

La evolución estructural de *Catsper1* en roedores está influida por la competición espermática, con efectos sobre la velocidad de natación del espermatozoide.

En la mayoría de los casos, la selección positiva de las proteínas reproductivas se ha identificado mediante análisis de divergencia de la secuencia a nivel de nucleótido. Sin embargo, otras clases de mutaciones, como por ejemplo inserciones y deleciones, también ocurren en las secuencias codificantes de las proteínas. Los cambios estructurales producidos por las inserciones y deleciones, que implican cambios en la longitud de las proteínas reproductivas, pueden estar sujetos a selección positiva, y por tanto bajo la influencia de fuerzas de selección sexual como la competición espermática.

El gen *Catsper1* codifica para una subunidad de Catsper, que es un canal de calcio expresado en el flagelo regulado por voltaje, siendo esencial para la motilidad espermática y la fecundación. La región N-terminal de *Catsper1* presenta un elevado número de inserciones y deleciones fijados por selección positiva. Sin embargo, aún no se conocen las fuerzas selectivas responsables de estos cambios, así como las implicaciones que pueden tener para la función del espermatozoide.

En este trabajo se evaluó si la fijación de cambios en longitud observados en el dominio intracelular de *Catsper1* son influidos por la competición espermática. Para ello, la secuencia de esta región fue analizada en un grupo de especies cercanas de roedores pertenecientes a la familia Muridae. Los resultados revelaron una correlación negativa entre la longitud de la secuencia de *Catsper1* y la masa relativa de testículo, que fue usado como índice de los niveles de competición espermática. Dado que *Catsper1* es importante para la motilidad del espermatozoide, también se analizó si la variación en longitud de la región N-terminal se asocia con cambios en la velocidad de natación del espermatozoide. Los resultados mostraron una correlación negativa entre la longitud de *Catsper1* y diversos parámetros de velocidad espermática.

4.4. Secuencias y análisis filogenéticos de *Catsper1*

Las secuencias codificantes del primer exón del gen *Catsper1* se obtuvieron para 16 especies de roedores de la familia Muridae y se alinearon. El alineamiento resultante abarcó un total de 1.880 nucleótidos, y reveló una alta divergencia de las secuencias así como un elevado número de inserciones y deleciones (desde este momento pasan a denominarse *indels*) (Figura suplementaria 1). Las secuencias de nucleótidos se tradujeron *in silico* a secuencias de aminoácidos. El

alineamiento de las secuencias de aminoácidos, que se corresponde con la región N-terminal de la proteína Catsper1, es mostrado en la Fig. 12.

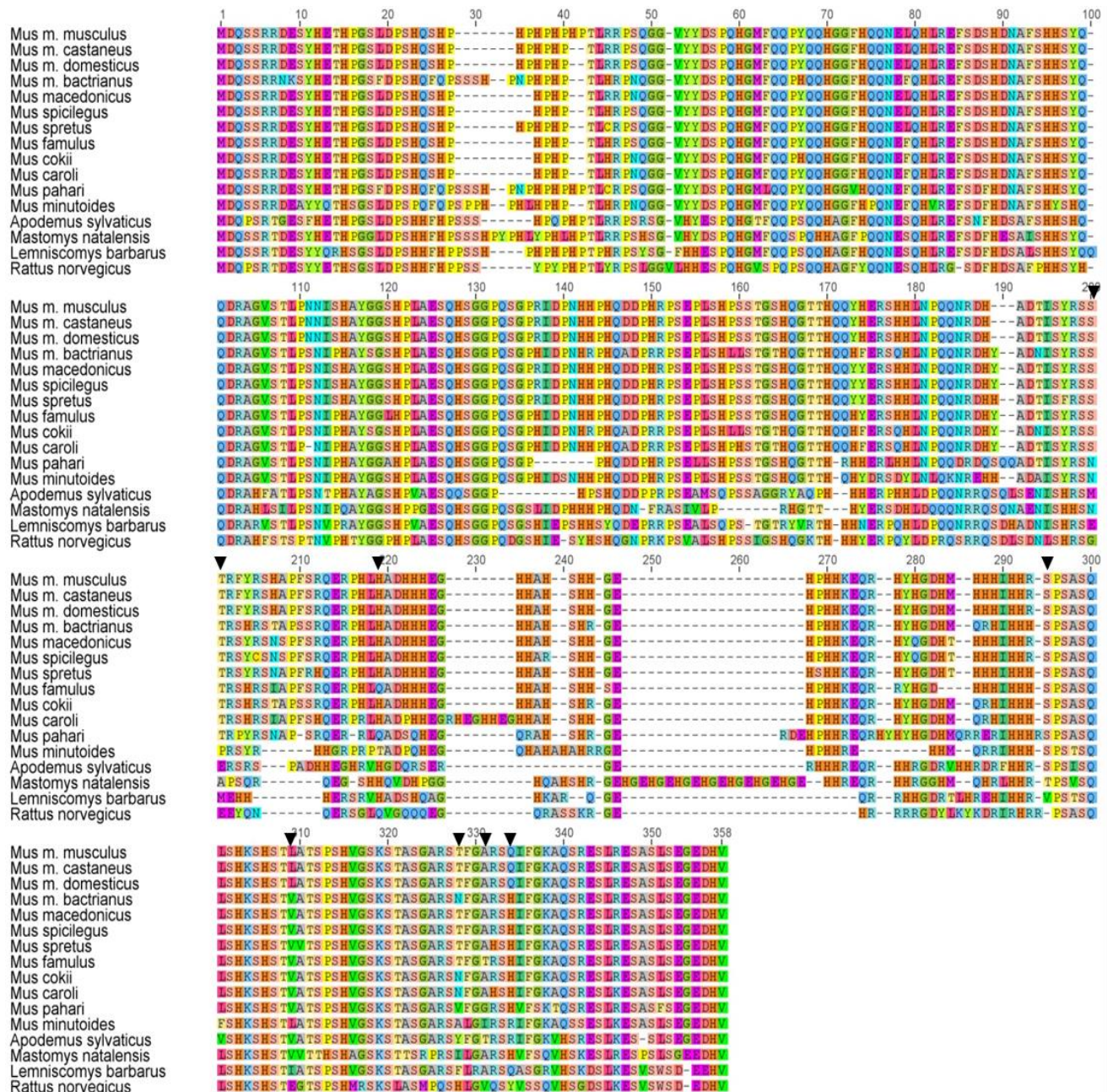


Figura 12. Alineamiento de la región N-terminal de Catsper1 en especies de roedores. Los guiones representan regiones sin alinear debido a la presencia de indels. Las flechas indican aminoácidos bajo selección positiva con una probabilidad posterior mayor a 0,95 de acuerdo con los modelos de selección (M2a y M8).

Partiendo del alineamiento de las secuencias de nucleótidos de *Catsper1*, se reconstruyeron filogenias de las especies analizadas por los métodos de Máxima Verosimilitud e Inferencia Bayesiana. Para ello, se utilizó el modelo de evolución GTR+G, con la distribución del parámetro gamma optimizada durante el análisis filogenético. Ambas filogenias mostraron una topología idéntica (Fig. 13). No obstante, estas filogenias mostraron algunas diferencias respecto a la filogenia establecida para las especies de roedores (Fig. 14). Esto sugiere que determinadas fuerzas selectivas pueden estar influyendo sobre *Catsper1*, alterando el patrón evolutivo esperado respecto a la relación filogenética de las especies.

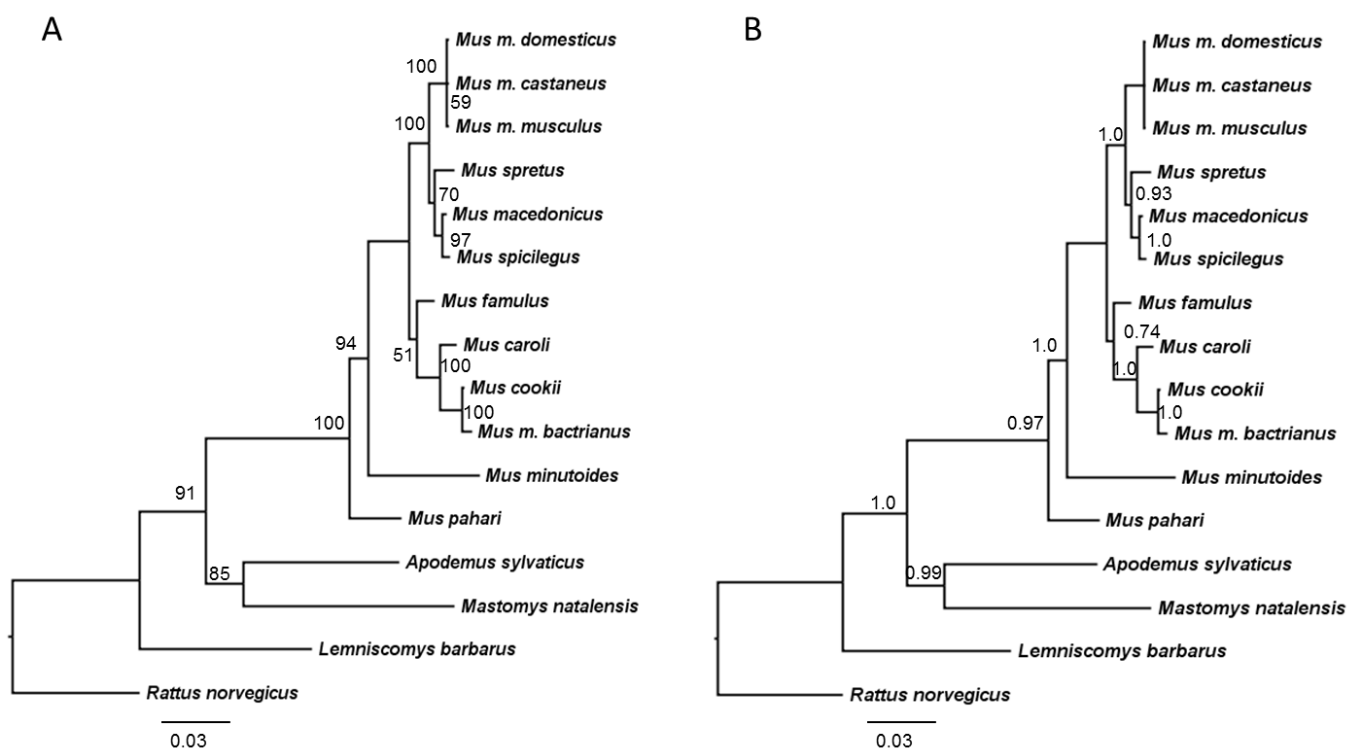


Figura 13. Árboles filogenéticos construidos por los métodos de A. Máxima Verosimilitud y B. Inferencia Bayesiana. El soporte de cada rama tras un análisis de bootstrap no paramétrico de 100 réplicas así como la probabilidad Bayesiana posterior se representan en los nodos internos de las filogenias de Máxima Verosimilitud e Inferencia Bayesiana respectivamente. Las longitudes de las ramas equivalen a las distancias genéticas, medidas en número de cambios sinónimos por codon. En el análisis Bayesiano, los parámetros se estimaron tras aplicar 100.000 generaciones y la frecuencia de muestreo fue cada 100 generaciones. Los árboles están representados a escala, la cual se indica en la parte inferior.

Utilizando el método de Máxima Parsimonia y considerando la posición de las especies en la filogenia, se infirieron un total de 58 indels (34 deleciones y 24 inserciones) del alineamiento de *Catsper1* (Fig. 14). 50 indels fueron únicos y congruentes con la topología del árbol, mientras que 8 indels fueron homoplásicos (es decir, se ha producido por convergencia evolutiva en linajes separados filogenéticamente). Dado que no se encontraron indels entre las secuencias de individuos de la misma especie, se consideró que las secuencias consenso eran representativas de cada linaje. Se identificaron un total de 35 indels en las ramas terminales de las especies, y fueron desde ninguna para *Mus cookii*, *Mus spretus*, *Mus spicilegus*, *Mus macedonicus*, *Mus m. domesticus* y *Mus m. castaneus*, hasta 9 para *Mastomys natalensis*.

Todos los indels identificados en el alineamiento presentaron longitudes de múltiplo de 3 nucleótidos (Fig. 14 y Figura suplementaria 1), por lo que el marco de lectura permaneció intacto en todas las secuencias. La tasa de indels en el primer exón de *Catsper1* fue significativamente mayor que el esperado respecto a la tasa calculada a nivel genómico, así como a la tasa calculada para el locus de *Catsper1* (ver metodología de Podlaha *et al.* 2005).

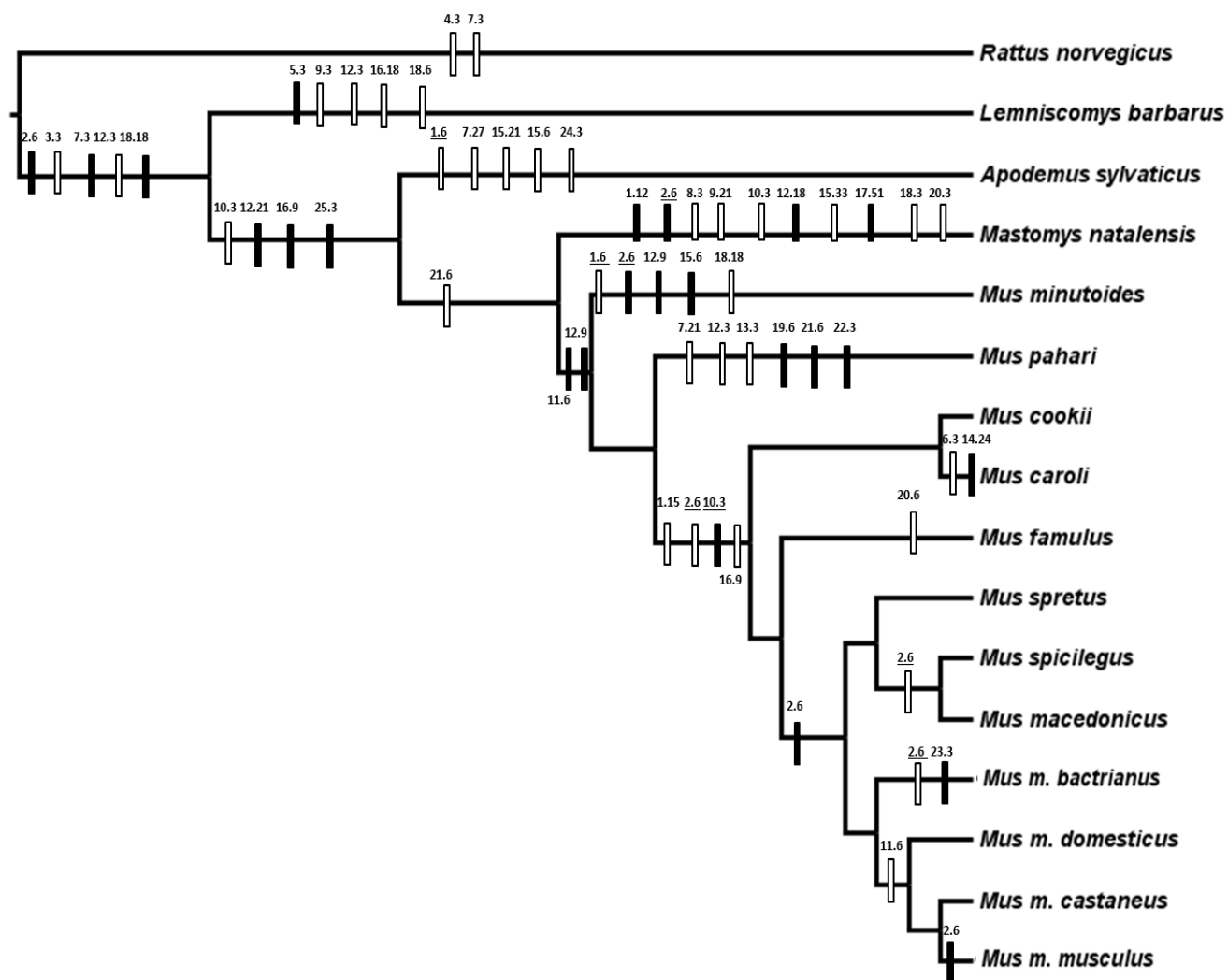


Figura 14. Cladograma representando la relación filogenética de las 16 especies de roedores analizadas. Los indels inferidos por parsimonia se encuentran mapeados a lo largo de la filogenia. Los rectángulos negros se corresponden con inserciones y los rectángulos blancos con deleciones. Los indels están anotados de acuerdo a la región de *Catsper1* donde han sido identificados (ver Figura suplementaria 1) seguidos por el número de nucleótidos que abarcan. Las anotaciones subrayadas indican indels homoplásicos. El árbol filogenético se construyó basándose en filogenias previamente publicadas de estas especies (Steppan *et al.*2004; Suzuki *et al.*2004; Martín-Coello *et al.*2009; Fabre *et al.*2012).

4.5. *Análisis de selección positiva en Catsper1*

Se llevaron a cabo una serie de análisis evolutivos con el fin de identificar aminoácidos bajo selección positiva en la secuencia de *Catsper1*. Una fuerte señal de selección positiva fue detectada en el primer exón de *Catsper1* cuando los modelos de selección (M2a y M8) fueron comparados con los correspondientes modelos neutrales (M1a, M7 y M8) (Tabla VIII). En ambos casos, los LRTs revelaron un mejor ajuste de los modelos de selección. Se detectaron un total de 10 y 9 aminoácidos bajo selección positiva bajo los modelos M2a y M8 respectivamente, con una

probabilidad posterior Bayesiana mayor que 0,95 (Tabla VIII, Fig. 12). 4 de estos 10 aminoácidos fueron residuos de histidina.

Tabla VIII. Análisis de selección positiva en Catsper1.

Modelo	Log-likelihood	LRT ^a	Estimas de parámetros ^b	Aminoácidos bajo selección positiva ^c
M1a	-2907,98	56,826**	$p_0 = 0,269$, $p_1 = 0,731$, $\omega_0 = 0,027$, $\omega_1 = 1$	No permitidos
M2a	-2879,57		$p_0 = 0,186$, $p_1 = 0,651$ $p_2 = 0,163$, $\omega_0 = 0$, $\omega_1 = 1$, $\omega_2 = 4,67$	124 (0,981*), 128 (0,954)*, 146 (0,979)*, 147(0,998**), 154 (0,991**), 165 (0,971*), 179 (0,985*), 198 (1,000**), 201 (0,993**), 204 (0,994**)
M7	-2908,01		$p = 0,041$, $q = 0,012$	No permitidos
M8	-2879,68	56,673**	$p_0 = 0,833$, $p = 0,0166$, $q = 0,0052$, $p_1 = 0,166$, $\omega = 4,535$	124 (0,991**), 146 (0,989*), 147 (0,998**), 154 (0,995**), 165 (0,983*), 179 (0,992**), 198 (1,000**), 201 (0,996**), 204 (0,997**)
M8a	-2907,98	56,613**	$p_0 = 0,269$, $p = 2,758$, $q = 99,0$, $p_1 = 0,731$, $\omega = 1,0$	No permitidos

Se alinearon 16 secuencias, analizándose un total de 235 codones.

L = longitud de las secuencias (en número de codones analizados).

^a LRTs significativos con un nivel de significancia del 99% (**).

^b Bajo el modelo M2a, los análisis Bayesianos empíricos (BEB) son capaces de clasificar los aminoácidos en tres clases de sitios de acuerdo a su valor posterior de ω . La clase p_0 incluye los aminoácidos bajo selección purificante ($0 < \omega < 1$), la clase p_1 los aminoácidos bajo evolución neutral con un ω aproximado a 1, y la clase p_2 aquellos aminoácidos estimados bajo selección positiva con un $\omega > 1$.

^c Aminoácidos bajo selección positiva con una probabilidad posterior $> 0,95$ (*) y $> 0,99$ (**) en un análisis Bayesiano Empírico (BEB).

4.6. Asociación entre la evolución de *Catsper1* y los niveles de competición espermática

Se aplicaron modelos evolutivos capaces de calcular tasas evolutivas promedio para cada linaje de una filogenia. Las tasas evolutivas (ω) estimadas para *Catsper1* fueron mayor que 1 para todas las especies de mridos excepto para *Lemniscomys barbarus* (Tabla suplementaria 4). Este resultado refleja la fuerte intensidad con la que la seleccin positiva acta sobre esta protena. Con el fin de evaluar la posible influencia de la competicin espermática sobre la tasa evolutiva de *Catsper1*, se correlacionaron los valores de ω especficos de cada especie con la masa testicular relativa a la masa corporal. Los anlisis de correlacin corregidos por efectos filogenéticos (PGLS) no mostraron una asociacin significativa entre estas dos variables (Tabla IX).

Tabla IX. Anlisis PGLS de *Catsper1* en relacin a la competicin espermática.

Variable dependiente	Variable predictora	n	Pendiente	R ²	F	p	λ^a	ES ^b	CL (-) ^b	CL (+) ^b
Omega (ω)	Masa corporal	16	0,002	0,539	13,939	0,003	0,999	0,053	-0,363	1,449
	Masa testicular		-0,092		0,14	0,714		0,056	-0,439	0,647
Longitud de secuencia	Masa corporal	15	0,149	0,778	8,895	0,011*	0	0,779	0,214	1,345
	Masa testicular		-21,159		33,055	0,001**		1,28	0,714	1,846

Los valores de $p < 0,05$ (*) y $p < 0,01$ (**) indican significancia estadística.

^a λ : parámetro de escala filogenética que indica la transformacin que hacen los datos respecto al modelo de evolucin Browniana para un mayor ajuste.

^bEfecto de tamao (ES) calculado a partir de los valores del estadístico F y sus lmites de confianza al 95% (CL). La significancia estadística se obtiene cuando los intervalos no incluyen al cero.

Por otro lado, se analizó si la variación estructural de la región N-terminal de Catsper1 se encuentra influida por la competición espermática. Para ello, se realizaron análisis de PGLS entre la longitud de la secuencia de esta región con la masa testicular residual. Debido a que el PGLS es un análisis de regresión basado en mínimos cuadrados, y estos son análisis sensibles a la presencia de outliers, se buscó entre los datos la existencia de outliers. Se encontró que la longitud de la secuencia de *Mastomys natalensis* era muy diferente de la del resto de especies (Figura suplementaria 2). Esto es principalmente debido a que la secuencia de esta especie posee una inserción específica de 51 pb, formada como consecuencia de una duplicación en tándem de 3 codones (Fig. 12 y Figura suplementaria 1). Por tanto, se descartó esta secuencia de los análisis de PGLS. Se encontró una correlación significativamente negativa entre la longitud de la región N-terminal de Catsper1 y la masa testicular residual (Tabla IX). La representación gráfica de la regresión entre el tamaño de la secuencia y la masa relativa de testículo se puede ver en la Fig. 15.

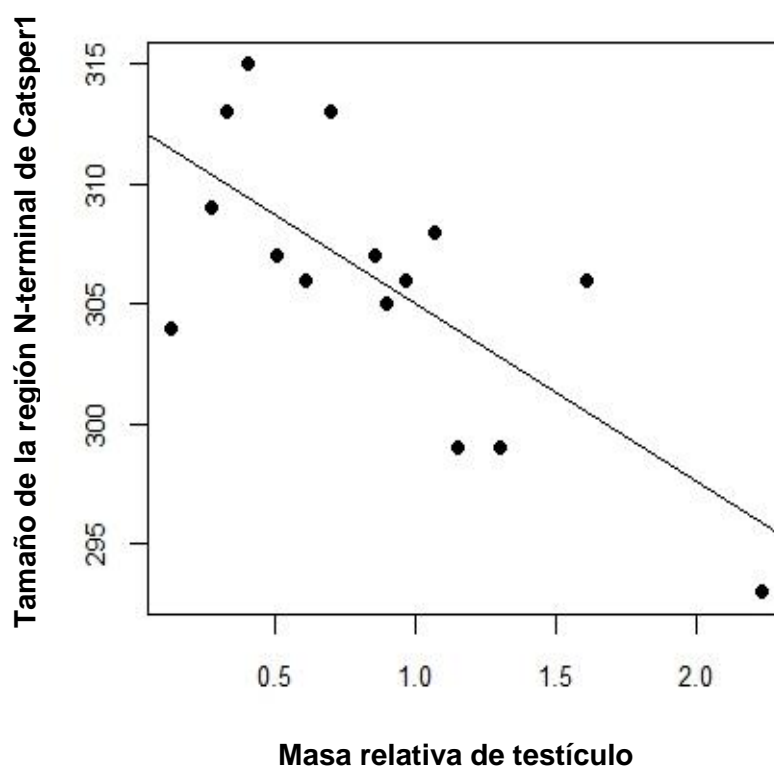


Figura 15. Relación entre los niveles de competición espermática (masa relativa de testículo) y la longitud de la secuencia de la región amino terminal de Catsper1.

4.7. *Asociación entre la evolución de *Catsper1* y la velocidad de natación espermática*

Dado el papel importante que tiene *Catsper1* en la motilidad espermática, se evaluó si la divergencia molecular y estructural del dominio N-terminal intracelular tiene efectos sobre las propiedades de natación del espermatozoide. No se obtuvieron asociaciones significativas cuando las tasas evolutivas de *Catsper1* fueron correlacionadas con parámetros de velocidad espermática, excepto para BCF (Tabla X). Por otra parte, sí se obtuvieron correlaciones significativamente negativas cuando la longitud de la región N-terminal fue analizada con los parámetros midiendo velocidad lineal (VSL, VAP, LIN, STR) (Figura suplementaria 3) y el parámetro de velocidad global obtenido del PCA (Tabla X, Fig. 16).

Tabla X. Análisis PGLS de Catsper1 en relación a los parámetros de velocidad espermática.

Variable dependiente	Variable predictora	n	Pendiente	R ²	F	p	λ^a	ES ^b	CL (-) ^b	CL (+) ^b
VCL	Omega (ω)	12	-19,59	0,171	2,062	0,182	0,689	0,439	-0,213	1,093
VAP	Omega (ω)	12	-9,384	0,021	0,214	0,811	0,786	0,146	-0,507	0,799
VSL	Omega (ω)	12	-8,237	0,012	0,12	0,736	0,885	0,109	-0,544	0,762
ALH	Omega (ω)	12	-0665	0,145	2,873	0,103	0	0,513	-0,140	1,166
BCF	Omega (ω)	12	0,056	0,688	25,32	0,001	0,51	1,244	0,591	1,898
LIN	Omega (ω)	12	-0,003	0,013	0,128	0,881	0,921	0,113	-0,540	0,766
STR	Omega (ω)	12	-2,873	-0,099	0,001	0,999	0,981	0,009	-0,643	0,663
Velocidad global	Omega (ω)	12	-1,009	0,015	0,137	0,874	0,849	0,123	-0,569	0,816
VCL	Longitud de secuencia	11	-0,975	0,512	1,966	0,191	0,754	0,430	-0,223	1,083
VAP	Longitud de secuencia	11	-2,003	0,164	8,340	0,016*	0,556	0,818	0,165	1,472
VSL	Longitud de secuencia	11	-2,375	0,216	10,48	0,009**	0,667	0,898	0,245	1,551
ALH	Longitud de secuencia	11	0,044	0,201	2,265	0,159	0,528	0,483	-0,21	1,176
LIN	Longitud de secuencia	11	-0,016	0,785	32,795	0,003**	0	1,402	0,709	2,095
STR	Longitud de secuencia	11	-0,009	0,626	15,06	0,001**	0,66	1,074	0,381	1,767
BCF	Longitud de secuencia	11	0,016	0,009	0,085	0,919	1	0,097	-0,595	0,789
Velocidad global	Longitud de secuencia	11	-0,195	0,475	10,03	0,005**	0,652	0,920	0,227	1,613

Los valores de $p < 0,05$ (*) y $p < 0,01$ (**) indican significancia estadística.

^aLos valores de λ indican el efecto filogenético.

^bEfecto de tamaño (ES) calculado a partir de los valores del estadístico F y sus límites de confianza al 95% (CL).

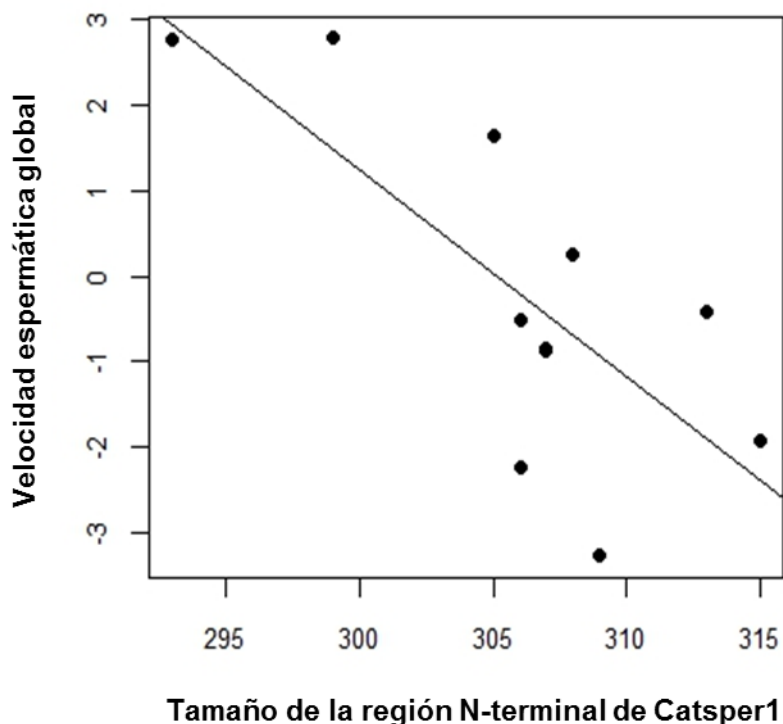


Figura 16. Correlación entre la longitud de secuencia de la región N-terminal de Catsper1 y la velocidad espermática.

4.8. Evolución de los residuos de histidina

Debido a su alto contenido en residuos de histidina, se ha planteado que el dominio N-terminal intracelular de Catsper1 está implicado en la regulación de la actividad del canal Catsper por pH. Por lo tanto, se analizó si la variación en el contenido de histidina observada entre las diferentes especies de roedores está relacionada con los niveles de competición espermática. Se calculó la proporción de histidinas en cada secuencia (Tabla suplementaria 4) y se llevó a cabo un PGLS utilizando la masa testicular residual como variable predictora. No se encontró una correlación significativa en este análisis ($F = 1,47$, $p = 0,25$) (Fig. 16). De un modo similar, no se observó una correlación significativa entre el porcentaje de histidinas y los parámetros de velocidad de natación ($F = 0,89$, $p = 0,42$). Tampoco se encontró una relación significativa de la cantidad de histidinas con la longitud de la secuencia ($F = 3,26$, $p = 0,094$), por lo que los cambios en el número de histidinas no son atribuibles a pérdidas o ganancias de secuencia por deleciones o inserciones.

Por otra parte, se detectaron dos ganancias globales de histidinas a través de la filogenia: la primera tiene lugar en el ancestro común de las especies de muridos después de la separación respecto a la rata (*Rattus norvegicus*), y la segunda ocurre entre *Mus pahari* y las especies restantes del género *Mus* (Fig.

17). Estas observaciones sugieren que la proporción de histidinas en Catsper1 evoluciona de acuerdo al patrón filogenético de las especies.

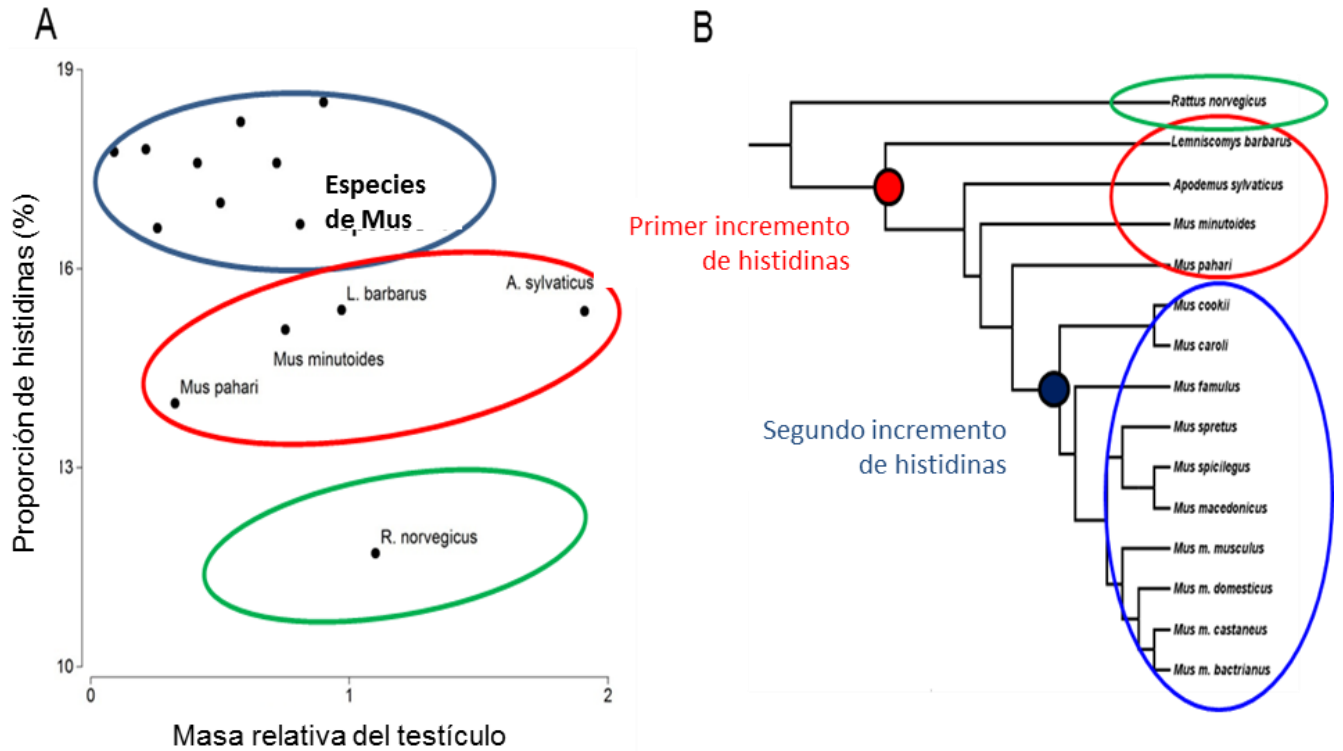


Figura 17. Evolución de las histidinas de Catsper1 en roedores. **A.** Diagrama de dispersión mostrando la distribución del porcentaje de histidinas, distinguiendo 3 subgrupos de especies. **B.** Cladograma en que se representan los momentos evolutivos en los que se han producido ganancias de histidinas.

BLOQUE III

Evolución de la proteína PKDREJ y el efecto de la competición espermática en roedores.

La evolución adaptativa de las proteínas reproductivas puede estar dirigida, al menos en parte, por fuerzas de selección sexual post-cópula como la competición espermática.

PKDREJ es una proteína que se expresa específicamente en el testículo y se localiza en la región anterior de la cabeza del espermatozoide. Se ha sugerido que PKDREJ en mamíferos es una proteína que interacciona con componentes de la zona pelúcida para desencadenar la reacción acrosómica. Estudios funcionales recientes en ratón han revelado que PKDREJ tiene efectos en el transporte del espermatozoide en el tracto reproductivo de la hembra y en el tiempo de la iniciación de la reacción acrosómica. Esto fue demostrado en análisis de competición espermática, donde espermatozoides de un ratón silvestre y un ratón transgénico competían por fecundar los mismos óvulos. Por tanto, estos hallazgos han hecho considerar a PKDREJ como un factor importante en la selección sexual post-cópula.

En este trabajo se analizó la evolución molecular del gen *Pkdrej* en dos familias de roedores pertenecientes a la superfamilia Muroidea: Muridae y Cricetidae. Los análisis evolutivos detectaron evidencia de selección positiva en el gen *Pkdrej* para especies de la familia Cricetidae, pero no para la familia Muridae. Posteriormente, se analizó si la divergencia de *Pkdrej* en las especies de cricétidos está asociada con los niveles de competición espermática. Se encontró que la tasa evolutiva de *Pkdrej* muestra una correlación significativa con la masa de los testículos respecto a la masa corporal en las especies de cricétidos. Además, los análisis evolutivos revelaron que la señal de selección positiva se concentra en aquellas especies con niveles más elevados de competición espermática.

Por otra parte, se analizó la evolución de PKDREJ en primates con el fin de observar si la influencia de la competición espermática sobre la evolución de esta proteína es un fenómeno generalizado en mamíferos.

4.9. *Secuencias y análisis filogenéticos de PKDREJ en roedores*

Para este estudio, se examinó la secuencia de la proteína PKDREJ en 25 especies de roedores pertenecientes a la superfamilia Muroidea. Doce de las especies pertenecen a la familia Muridae y 13 a la familia Cricetidae. Los fragmentos secuenciados de la secuencia codificante de *Pkdrej* tuvieron una longitud de entre 1.248 y 1.257 nucleótidos, y la traducción *in silico* de estas secuencias resultó en fragmentos de aminoácidos homólogos a las posiciones 213-629 de PKDREJ de *M. musculus* (Fig. 18). Este fragmento se corresponde con el dominio extracelular REJ de la proteína PKDREJ.

A

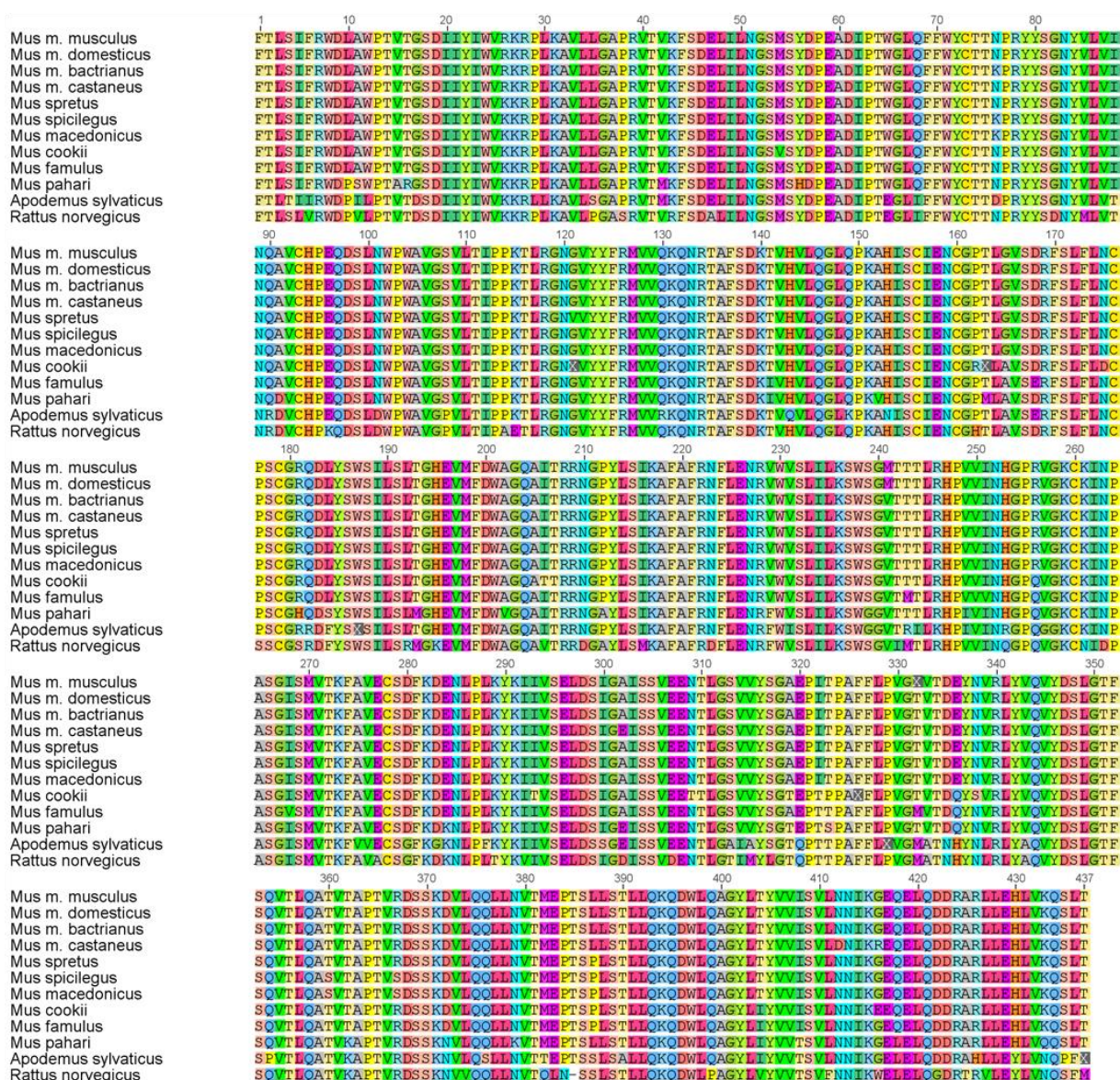


Figura 18. Alineamiento de aminoácidos de PKDREJ en roedores móridos (A) y cricétidos (B). Las flechas en el alineamiento de cricétidos indican aminoácidos bajo selección positiva con una probabilidad posterior mayor a 0,95 de acuerdo con los modelos de selección (M2a y M8).

B



Figura 18. (Continuación).

Tras obtener el alineamiento de las secuencias de nucleótidos de *Pkdrej* (Figura suplementaria 4), se reconstruyeron filogenias por los métodos de Máxima Verosimilitud e Inferencia Bayesiana, empleando el modelo de sustitución GTR+G+I tanto en múridos como en cricétidos. La proporción de sitios invariables fue fijada a 0,00002 en múridos, mientras que en cricétidos se fijó a 0,0858. En ambos casos, la distribución del parámetro gamma fue estimada. Las filogenias resultantes a partir de los dos métodos presentaron las mismas topologías tanto para las especies de múridos como de cricétidos (Fig. 19). Estas filogenias en general concuerdan con los árboles resueltos en estudios filogenéticos previos

(Steppan *et al.* 2004; Suzuki *et al.* 2004; Galewski *et al.* 2006; Neumann *et al.* 2006; Martín-Coello *et al.* 2009), aunque con algunas diferencias. En el caso de la familia Muridae, el clado que agrupa a las subespecies de *M. musculus* fue resuelto con un soporte bajo, dado que entre estas especies existe una politomía. Este fenómeno se ha observado con frecuencia en los análisis filogenéticos de roedores múridos (Sandstedt y Tucker 2005; Martín-Coello *et al.* 2009). La rama terminal de la especie *Mus cookii* se resolvió como adyacente al clado de las especies de *Mus* con una baja probabilidad, ya que se asume que es *Mus famulus* la especie más próxima a este género. En la familia Cricetidae, solo el clado agrupando *Microtus arvalis* con las especies del género *Pitimys* se resolvió con una probabilidad baja. Por lo tanto, para los análisis evolutivos se utilizaron topologías basadas en árboles filogenéticos establecidos de estudios previos (Figura suplementaria 5).

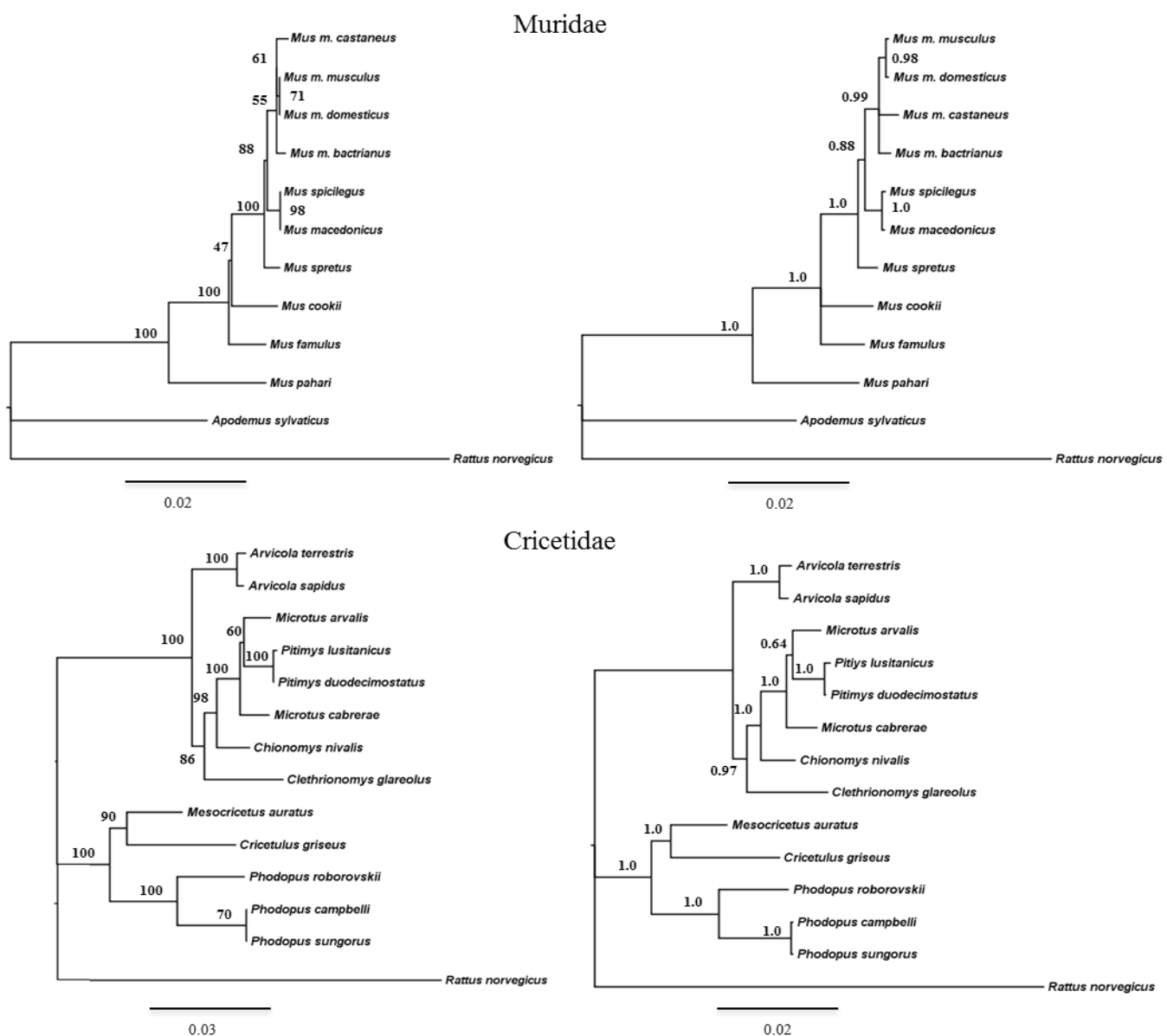


Figura 19. Árboles filogenéticos de *Pkdrej* en roedores. Los soportes de cada rama para un bootstrap no paramétrico de 100 réplicas y la probabilidad Bayesiana se representan en los nodos internos de las filogenias de Máxima Verosimilitud (Izquierda) e Inferencia Bayesiana (derecha). La longitud de las ramas indica las distancias genéticas. En los análisis Bayesianos se aplicaron 100.000 generaciones con una frecuencia de muestreo de 100 generaciones. Las escalas se indican en la parte inferior de cada árbol.

4.10. Evidencia de selección positiva en PKDREJ de roedores

Para detectar evidencia de selección positiva dentro de la proteína PKDREJ, se aplicaron modelos que tienen en cuenta la estimación de valores heterogéneos de ω entre los aminoácidos. Antes de realizar los análisis evolutivos para cada familia de roedores, se llevó a cabo un análisis preliminar incluyendo todas las especies de muroideos en una filogenia. Ambos modelos de selección (M2a y M8) ajustaron mejor los datos que los respectivos modelos neutrales (M1a, M7 y M8a) (Tabla XI). El modelo M8 detectó un aminoácido bajo selección positiva en la posición 47 ($P_{\text{BEB}} = 0,978$).

Para evaluar si la evolución adaptativa observada en PKDREJ es el resultado de selección positiva actuando en diferentes etapas de la evolución de los muroideos, los análisis de selección se aplicaron por separado a las dos familias de roedores estudiadas. En el caso de la familia Muridae, los modelos de selección (M2a y M8) fueron rechazados por el LRT en favor de los modelos neutrales (M1a, M7 y M8a), indicando que PKDREJ no experimenta evolución adaptativa en esta familia (Tabla XI). Los mismos resultados fueron obtenidos cuando se utilizaron las dos topologías posibles para resolver la politomía de las subespecies de *Mus*. En el caso de la familia Cricetidae, los LRTs mostraron evidencia de selección positiva (Tabla XI). Ambos modelos de selección (M2a y M8) también identificaron el aminoácido de la posición 47 sujeto a selección positiva dentro de esta familia, mientras que otro aminoácido en la posición 206 fue identificado bajo selección positiva únicamente por el modelo M8 (Tabla XI, Fig. 18B).

Tabla XI. Análisis de selección positiva en PKDREJ de roedores.

Especies	N	L	Modelos comparados	LRT ^a	Estimas de parámetros ^b	Aminoácidos bajo selección positiva ^c
Muroidea (superfamilia)	25	417	M1a-M2a	8,052 *	p0=0,646, p1=0,350, p2=0,003, ω 0=0,184, ω 1=1, ω 2=6,471	No detectados
			M7-M8	12,195 **	p0=0,982, p=0,680, q=0,919, p1=0,018,	47 (0,978*)
			M8a-M8	52,8845 **		
Muridae (familia)	12	437	M1a-M2a	0,179	p0=0,552, p1=0,448, ω 0 = 0,117, ω 1= 1	No permitidos
			M7-M8	0,381	p= 0,178, q= 0,168	No permitidos
			M8a-M8	0,292	p0=0,569, p= 14,512, q= 98,963, p1= 0,430, ω = 1	No permitidos
Cricetidae (familia)	13	426	M1a-M2a	9,376**	p0= 0,598, p1= 0,393, p2= 0,009, ω 0=0,123, ω 1=1, ω 2=6,551	47 (0,984*)
			M7-M8	11,167**	p0= 0,989, p= 0,265, q=0,312, p1= 0,011, ω = 5,679	47 (0,981*) 206 (0,957*)
			M8a-M8	8,242*		

N: número de secuencias.

L: longitud de las secuencias (número de codones).

^aLRTs significativos con un nivel de significancia del 95% (*), 99% (**). Se muestran los parámetros estimados para el modelo que mejor se ajusta.

^bBajo el modelo M2a, los análisis Bayesianos empíricos (BEB) son capaces de clasificar los aminoácidos en tres clases de sitios de acuerdo a su valor posterior de ω . La clase p_0 incluye los aminoácidos bajo selección purificante ($0 < \omega < 1$), la clase p_1 los aminoácidos bajo evolución neutral con un ω aproximado a 1, y la clase p_2 aquellos aminoácidos estimados bajo selección positiva con un $\omega > 1$.

^cAminoácidos bajo selección positiva con una probabilidad posterior $> 0,95$ (*) y $> 0,99$ (**) en un análisis Bayesiano Empírico (BEB).

4.11. Asociación entre la evolución de PKDREJ y los niveles de competición espermática en roedores

Con el fin de detectar si la competición espermática dirige la evolución de PKDREJ en roedores, se trató de encontrar asociaciones entre la tasa evolutiva de esta proteína y la masa relativa de testículo, utilizado como índice de competición espermática. Las estimas específicas de ω fueron menor que 1 para todos linajes, indicando que esta proteína se encuentra altamente conservada (Tabla suplementaria 5). Los análisis de regresión múltiple por PGLS revelaron una correlación significativamente positiva entre los valores de ω y la masa testicular corregida por la masa corporal (Tabla XII). La correlación entre la tasa evolutiva de PKDREJ en cricétidos y la masa relativa de testículo también fue significativa ($p = 0,003$) y se representa gráficamente en la Fig. 20. Los análisis de regresión para las especies de múridos no mostraron relaciones significativas (Tabla XII).

Tabla XII. Análisis PGLS de la tasa evolutiva de PKDREJ en relación a la competición espermática en roedores.

Familia	Variable predictora	n	Pendiente	R^2	F	p	λ^a	ES ^b	CL(-) ^b	CL(+) ^b
Muridae	Masa corporal	13	0,0002	0,058	0,256	0,628	0,584	0,188	-0,551	0,931
	Masa testicular		0,0737		0,917	0,370		0,340	-0,386	1,095
Cricetidae	Masa corporal	13	-0,0005	0,542	1,693	0,222	0,999	0,401	-0,219	1,020
	Masa testicular		0,0343		10,131	0,0097**		0,886	0,266	1,506

Los valores de ω estimados para cada linaje se utilizaron como variables dependientes.

El valor de $< 0,01$ (**) indica significancia estadística.

^a λ : parámetro de escala filogenética que indica la transformación que hacen los datos respecto al modelo de evolución Browniana para un mayor ajuste.

^bEfecto de tamaño (ES) calculado a partir de los valores del estadístico F y sus límites de confianza al 95% (CL).

La significancia estadística se obtiene cuando los intervalos no incluyen al cero.

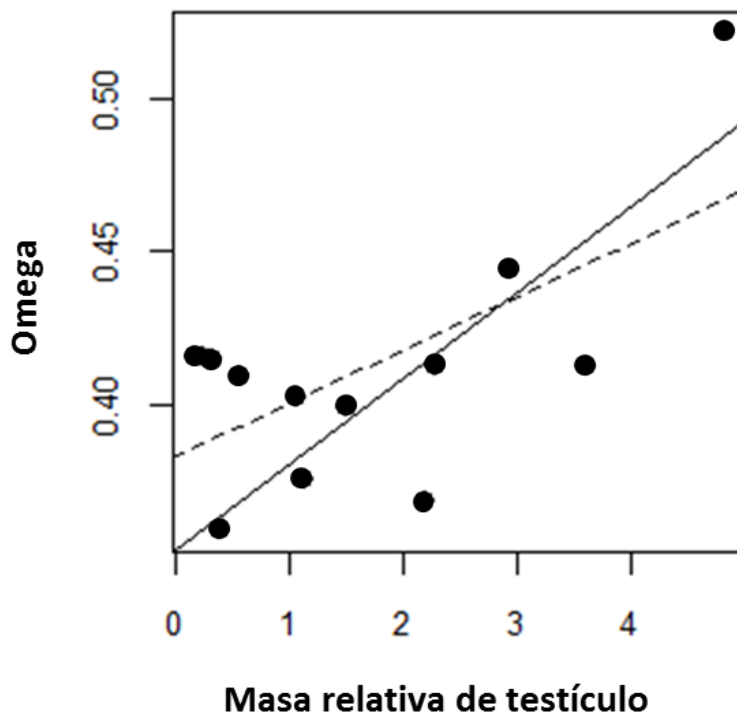


Figura 20. Relación entre la masa relativa de testículo y la tasa evolutiva de PKDREJ en cricétidos. La pendiente de la regresión estándar (línea discontinua) y la corregida por la filogenia (línea solida) son mostradas.

4.12. Detección de selección positiva en linajes de roedores

Con el fin de identificar señales de selección positiva sobre linajes específicos, se aplicaron modelos de evolución capaces de estimar tasas evolutivas entre codones dentro de cada secuencia de la filogenia. Estos análisis se realizaron por separado para las filogenias de Muridae y Cricetidae, tras observar que PKDREJ presenta un patrón evolutivo diferente entre estas familias de roedores. Dentro de la familia Cricetidae, se detectó evidencia de selección positiva en las ramas correspondientes a las especies *Cricetulus griseus* y *Mesocricetus auratus* (Tabla XIV). Estas dos especies presentan los niveles más altos de competición espermática. Dentro del linaje de *Cricetulus griseus* se detectó una señal de evolución adaptativa en el residuo 206. Los mismos análisis no detectaron evidencia de selección positiva para ningún linaje de la familia Muridae (Tabla XIII).

Tabla XIII. Análisis de selección positiva en especies de muridos.

Especie foreground	Probabilidad modelo alternativo	Probabilidad modelo nulo	LRT	p		Clases de codones ^a				Aminoácidos bajo selección positiva
						0	1	2a	2b	
<i>Apodemus sylvaticus</i>	-3090,832	-3091,084	0,504	0,478	proporción	0,438	0,336	0,128	0,098	No permitidos
					ω background	0,103	1,000	0,103	1,000	
					ω foreground	0,103	1,000	1,000	1,000	
<i>Mus pahari</i>	-3090,340	-3091,283	1,885	0,170	proporción	0,624	0,343	0,214	0,012	No permitidos
					ω background	0,179	1,000	0,179	1,000	
					ω foreground	0,179	1,000	1,000	1,000	
<i>Mus cookii</i>	-3087,932	-3088,587	1,310	0,252	proporción	0,279	0,207	0,295	0,219	No permitidos
					ω background	0,047	1,000	0,047	1,000	
					ω foreground	0,047	1,000	1,000	1,000	
<i>Mus famulus</i>	-3088,449	-3089,885	2,872	0,090	proporción	0,356	0,249	0,233	0,162	No permitidos
					ω background	0,084	1,000	0,084	1,000	
					ω foreground	0,150	1,000	1,000	1,000	
<i>Mus spretus</i>	-3088,353	-3090,257	3,808	0,051	proporción	0,382	0,280	0,195	0,143	No permitidos
					ω background	0,078	1,000	0,078	1,000	
					ω foreground	0,078	1,000	1,000	1,000	
<i>Mus spicilegus</i>	-3087,474	-3088,769	2,590	0,108	proporción	0,289	0,201	0,300	0,210	No permitidos
					ω background	0,063	1,000	0,063	1,000	
					ω foreground	0,063	1,000	1,000	1,000	
<i>Mus macedonicus</i>	-3087,474	-3088,769	2,590	0,108	proporción	0,289	0,201	0,300	0,210	No permitidos
					ω background	0,063	1,000	0,063	1,000	
					ω foreground	0,063	1,000	1,000	1,000	

Tabla XIII. Continuación

<i>Mus m. domesticus</i>					proporción	0,615	0,370	0,009	0,006	
	-3088,278	-3089,568	2,580	0,108	ω background	0,175	1,000	0,175	1,000	No permitidos
					ω foreground	0,175	1,000	1,000	1,000	
<i>Mus m. musculus</i>					proporción	0,615	0,370	0,009	0,006	
	-3088,278	-3089,568	4,789	0,108	ω background	0,175	1,000	0,175	1,000	No permitidos
					ω foreground	0,175	1,000	1,000	1,000	
<i>Mus m. castaneus</i>					proporción	0,608	0,380	0,008	0,005	
	-3088,963	-3089,575	1,224	0,269	ω background	0,171	1,000	0,171	1,000	No permitidos
					ω foreground	0,171	1,000	1,000	1,000	
<i>Mus m. bactrianus</i>					proporción	0,614	0,372	0,009	0,006	
	-3088,562	-3089,782	2,440	0,118	ω background	0,174	1,000	0,174	1,000	No permitidos
					ω foreground	0,174	1,000	1,000	1,000	

^aSe indican la proporción y los valores de ω estimados para cada clase de codones. Clase 0: codones bajo selección positiva ($0 < \omega < 1$) tanto en el background como en el foreground; Clase 1: codones evolucionando neutralmente ($\omega = 1$) tanto en el background como en el foreground; Clases 2a: proporción de codones bajo selección positiva en el foreground y bajo selección purificante en el background; Clase 2b: proporción de codones bajo selección positiva en el foreground y bajo evolución neutral en el background. En aquellos linajes donde los modelos neutrales fueron aceptados contra el modelo alterativo, los codones del foreground en las clases 2a y 2b se consideran bajo evolución neutral.

Tabla XIV. Análisis de selección positiva en especies de cricétidos.

Especie <i>foreground</i>	Probabilidad modelo alternativo	Probabilidad modelo nulo	LRT ^a	p		Clases de codones ^b				Aminoácidos bajo selección positiva
						0	1	2a	2b	
<i>Arvicola sapidus</i>	-4711,245	-4711,245	0,000	1,000	proporción	0,608	0,392	0,000	0,000	No permitidos
					ω background	0,120	1,000	0,120	1,000	
					ω foreground	0,120	1,000	1,000	1,000	
<i>Arvicola terrestris</i>	-4711,245	-4711,245	0,000	1,000	proporción	0,608	0,392	0,000	0,000	No permitidos
					ω background	0,120	1,000	0,120	1,000	
					ω foreground	0,120	1,000	1,000	1,000	
<i>Clethrionomys glareolus</i>	-4711,138	-4711,138	0,000	1,000	proporción	0,575	0,374	0,031	0,020	No permitidos
					ω background	0,112	1,000	0,112	1,000	
					ω foreground	0,112	1,000	1,000	1,000	
<i>Chionomys nivalis</i>	-4710,823	-4710,823	0,000	1,000	proporción	0,546	0,353	0,061	0,040	No permitidos
					ω background	0,108	1,000	0,108	1,000	
					ω foreground	0,108	1,000	1,000	1,000	
<i>Microtus arvalis</i>	-4711,245	-4711,245	0,000	1,000	proporción	0,608	0,392	0,000	0,000	No permitidos
					ω background	0,120	1,000	0,120	1,000	
					ω foreground	0,120	1,000	1,000	1,000	
<i>Microtus cabreræ</i>	-4711,245	-4711,245	0,000	1,000	proporción	0,608	0,392	0,000	0,000	No permitidos
					ω background	0,120	1,000	0,120	1,000	
					ω foreground	0,120	1,000	1,000	1,000	
<i>Pitimys duodecimostatus</i>	-4711,245	-4711,245	0,000	1,000	proporción	0,608	0,392	0,000	0,000	No permitidos
					ω background	0,120	1,000	0,120	1,000	
					ω foreground	0,120	1,000	1,000	1,000	

Tabla XIV Continuación

<i>Pitimys lusitanicus</i>					proporción	0,608	0,392	0,000	0,000	
	-4711,245	-4711,245	0,000	1,000	ω background	0,120	1,000	0,120	1,000	No permitidos
					ω foreground	0,120	1,000	1,000	1,000	
<i>Mesocricetus auratus</i>					proporción	0,597	0,379	0,015	0,009	
	-4710,864	-4711,208	4,789	0,029*	ω background	0,117	1,000	0,117	1,000	
					ω foreground	0,117	1,000	1,000	1,000	
<i>Cricetulus griseus</i>					proporción	0,599	0,390	0,006	0,004	
	-4706,814	-4711,218	8,810	0,003**	ω background	0,124	1,000	0,124	1,000	206R (0,988*)
					ω foreground	0,124	1,000	26,414	26,414	
<i>Phodopus sungorus</i>					proporción	0,600	0,385	0,009	0,006	
	-4710,910	-4711,231	0,644	0,422	ω background	0,119	1,000	0,119	1,000	No permitidos
					ω foreground	0,119	1,000	5,832	5,832	
<i>Phodopus campbelli</i>					proporción	0,600	0,385	0,009	0,006	
	-4710,910	-4711,231	0,644	0,422	ω background	0,119	1,000	0,119	1,000	No permitidos
					ω foreground	0,119	1,000	1,000	1,000	
<i>Phodopus roborovskii</i>					proporción	0,608	0,392	0,000	0,000	
	-4711,245	-4711,245	0,000	1,000	ω background	0,120	1,000	0,120	1,000	No permitidos
					ω foreground	0,120	1,000	1,000	1,000	

^aEl modelo alternativo se aceptó frente al modelo nulo cuando el LRT fue significativo frente a una distribución chi-cuadrado nula asintótica con un valor crítico de 2,71 en un nivel de significancia del 95%. Los LRT significativos con una $p < 0,05$ (*) y $< 0,01$ (**) son indicados.

^bSe indican la proporción y los valores de ω estimados para cada clase de codones. Clase 0: codones bajo selección positiva ($0 < \omega < 1$) tanto en el background como en el foreground; Clase 1: codones evolucionando neutralmente ($\omega = 1$) tanto en el background como en el foreground; Clase 2a: proporción de codones bajo selección positiva en el foreground y bajo selección purificante en el background; Clase 2b: proporción de codones bajo selección positiva en el foreground y bajo evolución neutral en el background. En aquellos linajes donde los modelos neutrales fueron aceptados contra el modelo alterativo, los codones del foreground en las clases 2a y 2b se consideran bajo evolución neutral. Los aminoácidos sujetos a selección positiva se muestran con la probabilidad posterior Bayesiana entre paréntesis.

4.13. Secuencias y análisis filogenéticos de *PKDREJ* en primates

Se tomaron las secuencias de *PKDREJ* para un grupo de especies de primates, obtenidas en un estudio interespecífico que se había realizado previamente (Hamm *et al.* 2007). Se reconstruyeron filogenias de primates a partir de un alineamiento de las secuencias de nucleótidos de *PKDREJ*, utilizando para ello los métodos de Máxima verosimilitud e Inferencia Bayesiana. El modelo de sustitución empleado fue GTR+G y la distribución del parámetro gama fue optimizada. Ambas metodologías produjeron un árbol filogenético similar al obtenido por Hamm *et al.* (2007) utilizando la metodología de Neighbor-Joining (Fig. 21). Esta topología está a su vez en concordancia con la filogenia establecida para primates (Goodman *et al.* 1998). Por tanto, esta fue la filogenia que se utilizó para los análisis evolutivos de *PKDREJ* en primates.

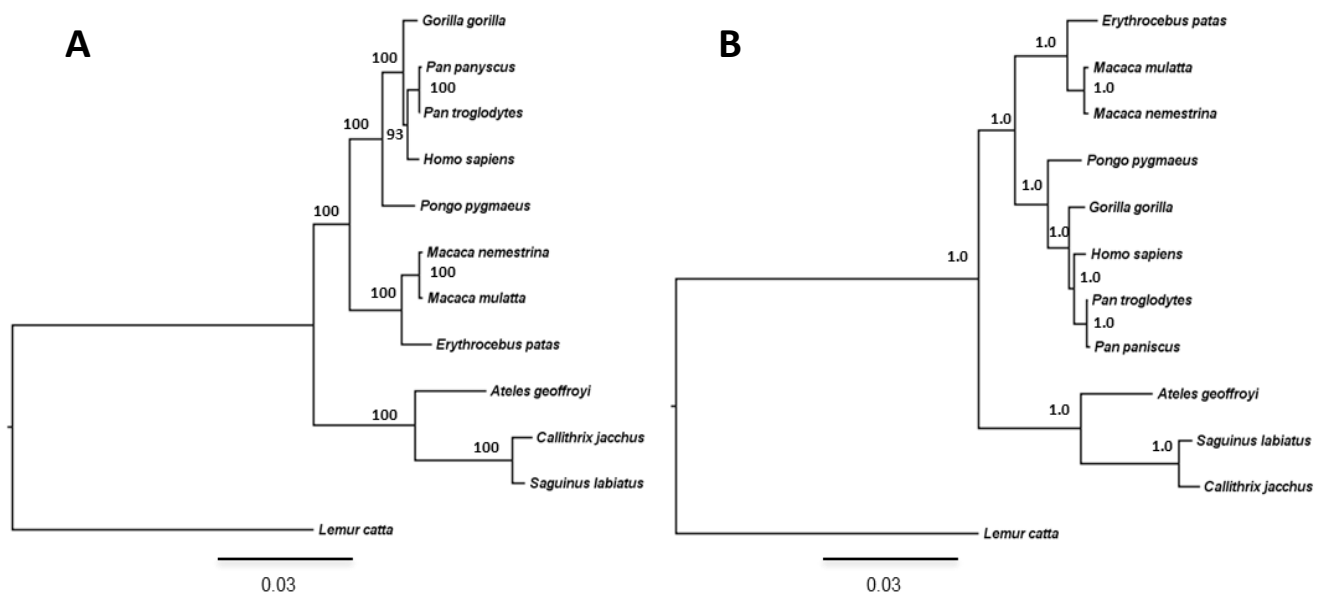


Figura 21. Árboles filogenéticos de *PKDREJ* en primates. EL soporte de cada rama tras un análisis de bootstrap no paramétrico de 100 réplicas así como la probabilidad Bayesiana posterior se representan en los nodos internos de las filogenias de Máxima Verosimilitud (A) e Inferencia Bayesiana (B) respectivamente. La longitud de las ramas indican las distancias genéticas, estimadas como el número de cambios sinónimos por codon. En el análisis Bayesiano, los parámetros se estimaron tras aplicar 50.000 generaciones y la frecuencia de muestreo fue de cada 100 generaciones. Las escalas se indican en la parte inferior de cada árbol.

4.14. Asociación entre la evolución de PKDREJ y los niveles de competición espermática en primates.

Aprovechando que se analizaron secuencias de PKDREJ en primates, se examinó si la competición espermática influye sobre la evolución esta proteína también en este grupo de mamíferos. Los análisis de regresión corregidos por efectos filogenéticos no revelaron asociaciones significativas entre los valores de ω estimados para cada especie y los valores de masa testicular corregida por la masa corporal (Tabla suplementaria 6), bien para la secuencia completa de PKDREJ o bien para la región correspondiente al dominio extracelular REJ (la secuencia analizada en roedores) (Tabla XV).

Tabla XV. Análisis PGLS de la tasa evolutiva de PKDREJ en relación a la competición espermática en primates.

Familia	Variable predictora	n	Pendiente	R^2	F	p	λ^a	ES ^b	CL(-) ^b	CL(+) ^b
PKDREJ completa	Log-Masa corporal	12	-0,009	0,089	0,237	0,638	0,999	0,161	-0,492	0,815
	Log-Masa testicular		0,0081		0,644	0,443		0,264	-0,389	0,918
Dominio REJ	Log-Masa corporal	12	-0,041	0,253	0,039	0,848	0,999	0,065	-0,587	-0,719
	Log-Masa testicular		0,014		3,006	0,117		0,549	-0,103	1,203

Los valores de ω estimados para cada linaje se utilizaron como variables dependientes.

El valor de $< 0,01$ (**) indica significancia estadística.

^a λ : parámetro de escala filogenética que indica la transformación que hacen los datos respecto al modelo de evolución Browniana para un mayor ajuste.

^bEfecto de tamaño (ES) calculado a partir de los valores del estadístico F y sus límites de confianza al 95% (CL). La significancia estadística se obtiene cuando los intervalos no incluyen al cero.

4.15. Detección de selección positiva en linajes de primates

Los análisis de selección positiva por linajes no detectaron evidencia de evolución adaptativa para ninguna especie de primate, bien para la secuencia completa de PKDREJ o para la secuencia del dominio REJ (Tabla XVI).

Tabla XVI. Análisis de selección positiva en especies de primates

Especie <i>foreground</i>	PKDREJ completa				Dominio REJ			
	Probabilidad del modelo alternativo	Probabilidad del modelo nulo	LRT	p	Probabilidad del modelo alternativo	Probabilidad del modelo nulo	LRT	p
<i>Gorilla gorilla</i>	-14225,85	-14225,85	0,000	1,000	-14225,85	-14225,85	0,000	1,000
<i>Homo sapiens</i>	-14225,87	-14225,87	0,000	1,000	-14225,87	-14225,87	0,000	1,000
<i>Pan troglodytes</i>	-14225,94	-14225,94	0,000	0,987	-14225,94	-14225,94	0,000	0,987
<i>Pan paniscus</i>	-14225,94	-14225,97	0,060	0,807	-14225,94	-14225,97	0,060	0,807
<i>Pongo pygmaeus</i>	-14226,06	-14226,06	0,003	0,955	-14226,06	-14226,06	0,003	0,955
<i>Erythrocebus patas</i>	-14223,52	-14223,52	0,000	1,000	-14223,52	-14223,52	0,000	1,000
<i>Macaca mulatta</i>	-14226,06	-14226,06	0,000	1,000	-14226,06	-14226,06	0,000	1,000
<i>Macaca nemestrina</i>	-14226,06	-14226,06	0,000	0,983	-14226,06	-14226,06	0,000	0,983
<i>Ateles geoffroyi</i>	-14223,52	-14223,52	0,000	0,986	-14223,52	-14223,52	0,000	0,986
<i>Saguinus labiatus</i>	-14224,30	-14224,40	0,190	0,663	-14224,30	-14224,40	0,190	0,663
<i>Callithrix jacchus</i>	-14223,72	-14224,40	1,358	0,244	-14223,72	-14224,40	1,358	0,244
<i>Lemur catta</i>	-14226,06	-14226,06	0,000	1,000	-14226,06	-14226,06	0,000	1,000

BLOQUE IV

Coevolución entre dos proteínas implicadas en la fusión de gametos: Izumo1 y Cd9.

Se ha observado que una alta proporción de proteínas implicadas en la reproducción sexual evolucionan rápidamente, a menudo como resultado de la selección positiva (o selección Darwiniana). No obstante, el conocimiento sobre las fuerzas selectivas que dirigen los cambios adaptativos en las proteínas reproductivas es escaso. Los procesos de coevolución entre proteínas de los gametos masculino y femenino implicadas en la fecundación puede ser una causa potencial de su rápida divergencia.

Hasta ahora, solamente se han identificado dos proteínas que son esenciales para la fusión del espermatozoide y el óvulo en ratón: Izumo1 en el espermatozoide y Cd9 en el óvulo. Aunque el papel de estas proteínas en la fecundación no ha sido completamente dilucidado, se ha sugerido que bien podrían actuar como fusógenos interaccionando en *trans* con proteínas en la superficie de otras células, o bien como reguladores de otros fusógenos a través de interacciones en *cis*.

En este capítulo se analizó la evolución de Izumo1 y Cd9 en un grupo de roedores. Para evaluar una posible interacción entre Izumo1 y Cd9, examinamos su grado de coevolución basándonos en los análisis de correlación de sus tasas evolutivas.

Los análisis evolutivos mostraron evidencia de que ambas proteínas evolucionan adaptativamente, siendo la selección positiva más intensa sobre Izumo1. Concretamente, la señal de selección positiva en ambas proteínas se concentró en dominios extracelulares. Por otra parte, la observación de una señal de coevolución entre Izumo1 y Cd9 sugiere la existencia de algún tipo de interacción entre estas dos proteínas, aunque no hay evidencia de que Izumo1 y Cd9 interaccionen físicamente como fusógenos.

La divergencia adaptativa de Izumo1 y Cd9 podría influir sobre la compatibilidad reproductiva entre las especies y, por tanto, estas proteínas podrían tener un papel importante en los sistemas de reconocimiento específico de espermatozoide-óvulo.

4.16. Secuencias y análisis filogenéticos

Para desarrollar los análisis evolutivos de las proteínas de estudio, se obtuvieron las secuencias codificantes de los genes *Izumo1* y *Cd9* para nueve especies de muridos. El uso de este grupo de especies filogenéticamente próximas favoreció la construcción de alineamientos robustos para cualquier combinación de parámetros en ambos genes. El alineamiento de las secuencias de *Izumo1* tuvo una longitud total de 1.218 nucleótidos y se identificaron un total de 6 regiones con inserciones y deleciones (indels) (Figura suplementaria 6). Tras la traducción *in silico* de las secuencias de nucleótidos, se obtuvo la secuencia de aminoácidos representada en la Fig. 22. Se calculó un valor

promedio de ω de 0.937 a lo largo de todas las secuencias y especies, y una porción de todas las comparaciones entre pares de especies mostró una relación d_N/d_S mayor que 1 (Fig. 24). El gen *Cd9* mostró un mayor grado de conservación que *Izumo1*, reflejado por una ausencia de indels en su alineamiento (Fig. 23, Figura suplementaria 7), un valor promedio de ω menor (0.334) y una distribución de todos los cocientes d_N/d_S estimados de comparaciones de a pares menor que 1 (Fig. 24).

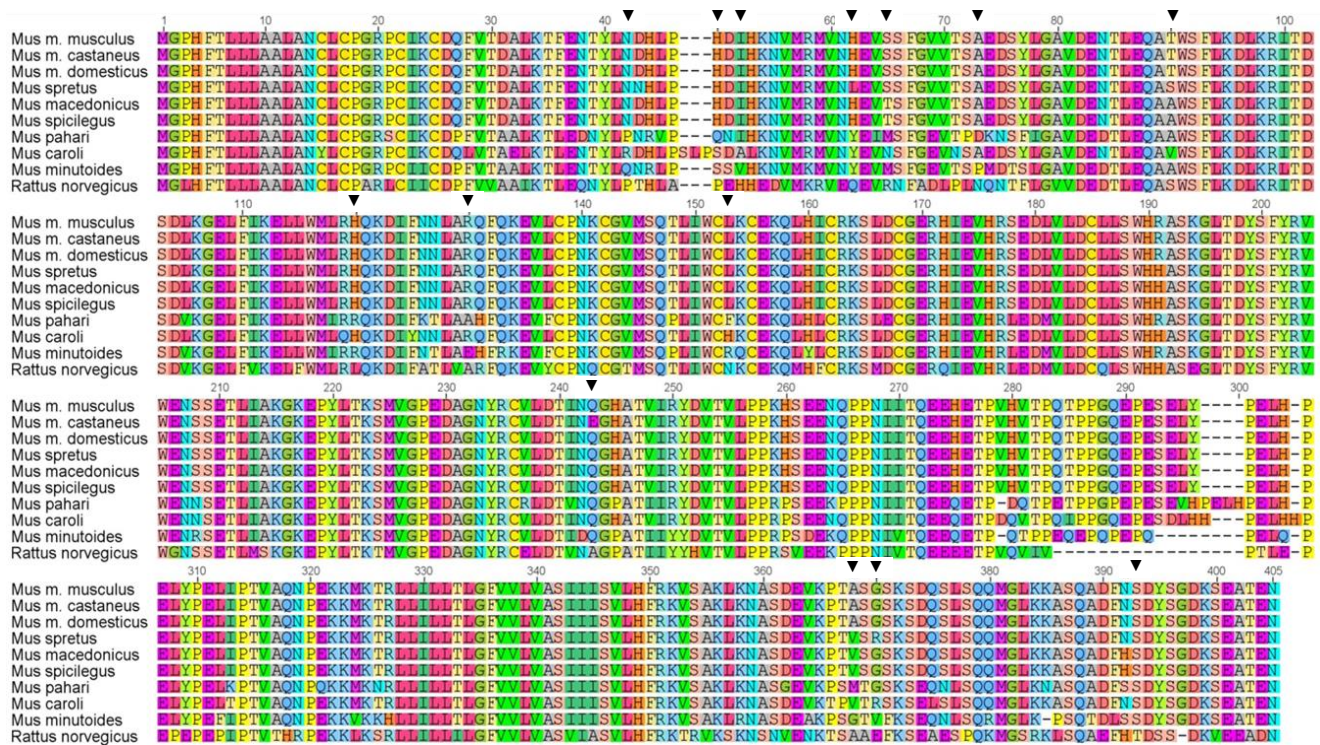


Figura 22. Alineamiento de aminoácidos de *Izumo1* en roedores múridos. Las flechas indican aminoácidos bajo selección positiva con una probabilidad posterior mayor a 0,95 de acuerdo con los modelos de selección (M2a y M8).

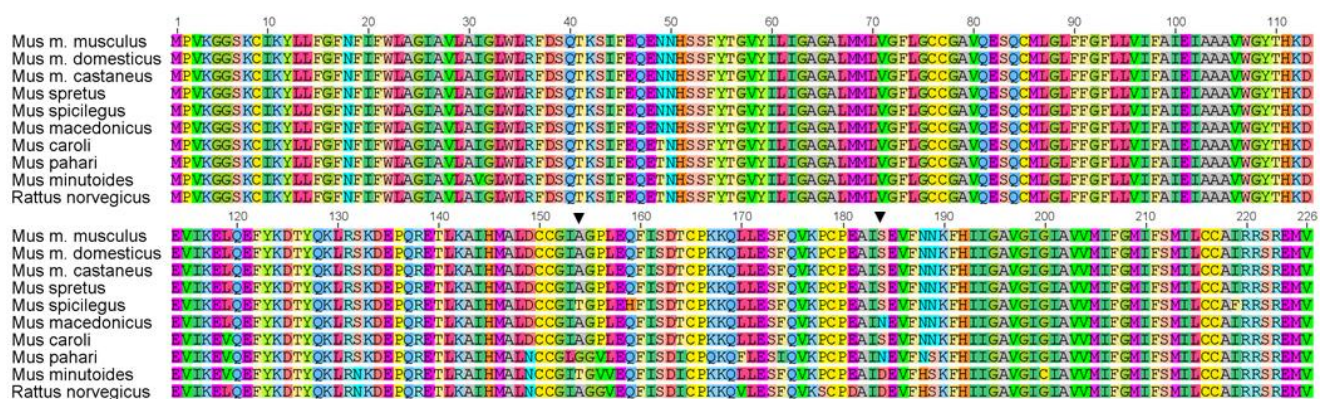


Figura 23. Alineamiento de aminoácidos de *Cd9* en roedores múridos. Las flechas indican aminoácidos bajo selección positiva con una probabilidad posterior mayor a 0,95 de acuerdo con los modelos de selección (M2a y M8).

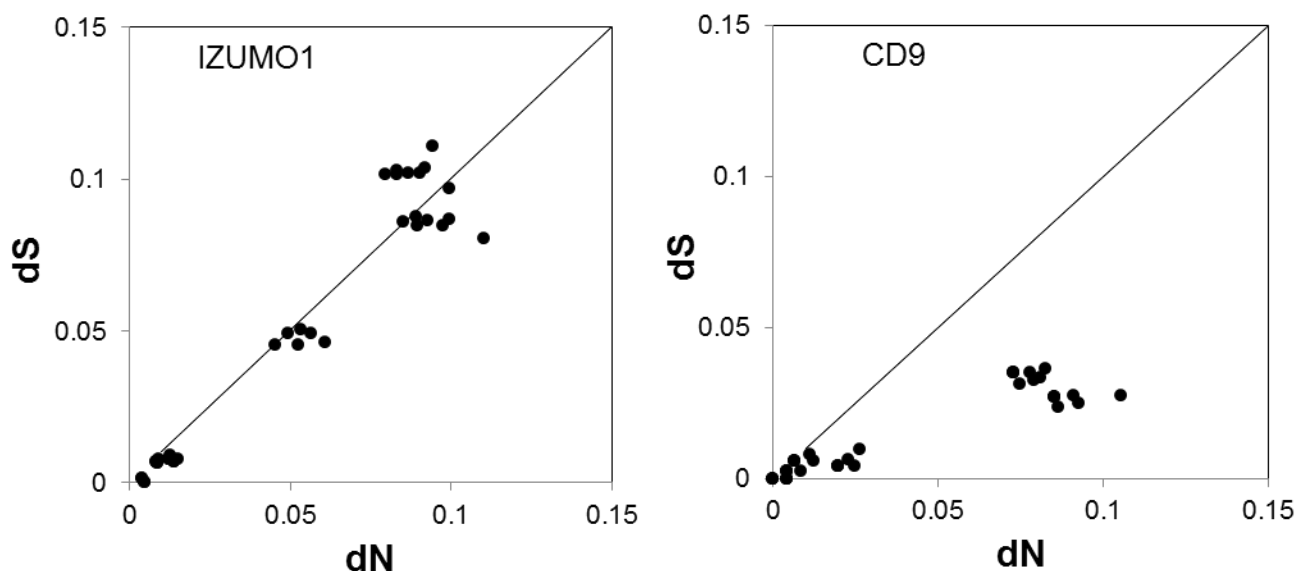


Figura 24. Distribución de tasas evolutivas de Izumo1 y Cd9. Cada punto representa el valor de ω (dN/ds) estimado entre pares de especies. La línea diagonal indica cuando los valores de dN y ds son iguales, dando un valor de $\omega = 1$.

El modelo de sustitución de nucleótidos que mejor describe la evolución de los genes *Izumo1* y *Cd9* fue GTR+G, estimando la distribución del parámetro gamma para la reconstrucción de las filogenias. Se observaron algunas diferencias entre las topologías construidas por Máxima Verosimilitud y por Inferencia Bayesiana tanto para las filogenias de *Izumo1* como las de *Cd9* (Fig. 25 y 26). Además, ninguna de los árboles generados mostró una topología de acuerdo con la filogenia establecida para las especies analizadas (Suzuki *et al.* 2004; Fabre *et al.* 2012). En la filogenia de *Izumo1* se produjo una politomía entre las subespecies de *M. musculus* (*M. m. musculus* y *M. m. castaneus*) que no se pudo resolver con un soporte significativo por el método Bayesiano (Sandstedt y Tucker 2005; Martín-Coello *et al.* 2009)

En el caso de *Cd9*, ambas metodologías solo fueron capaces de resolver con soporte las ramificaciones de *Mus caroli*, *Mus pahari* y *Mus minutoides*. La discordancia entre las filogenias inferidas con los genes de estudio y la establecida para las especies sugiere que tanto *Izumo1* como *Cd9* experimentan trayectorias evolutivas diferentes a la impuesta por la relación filogenética de las especies, posiblemente debido a la acción de fuerzas selectivas que dirigen la evolución de estos genes.

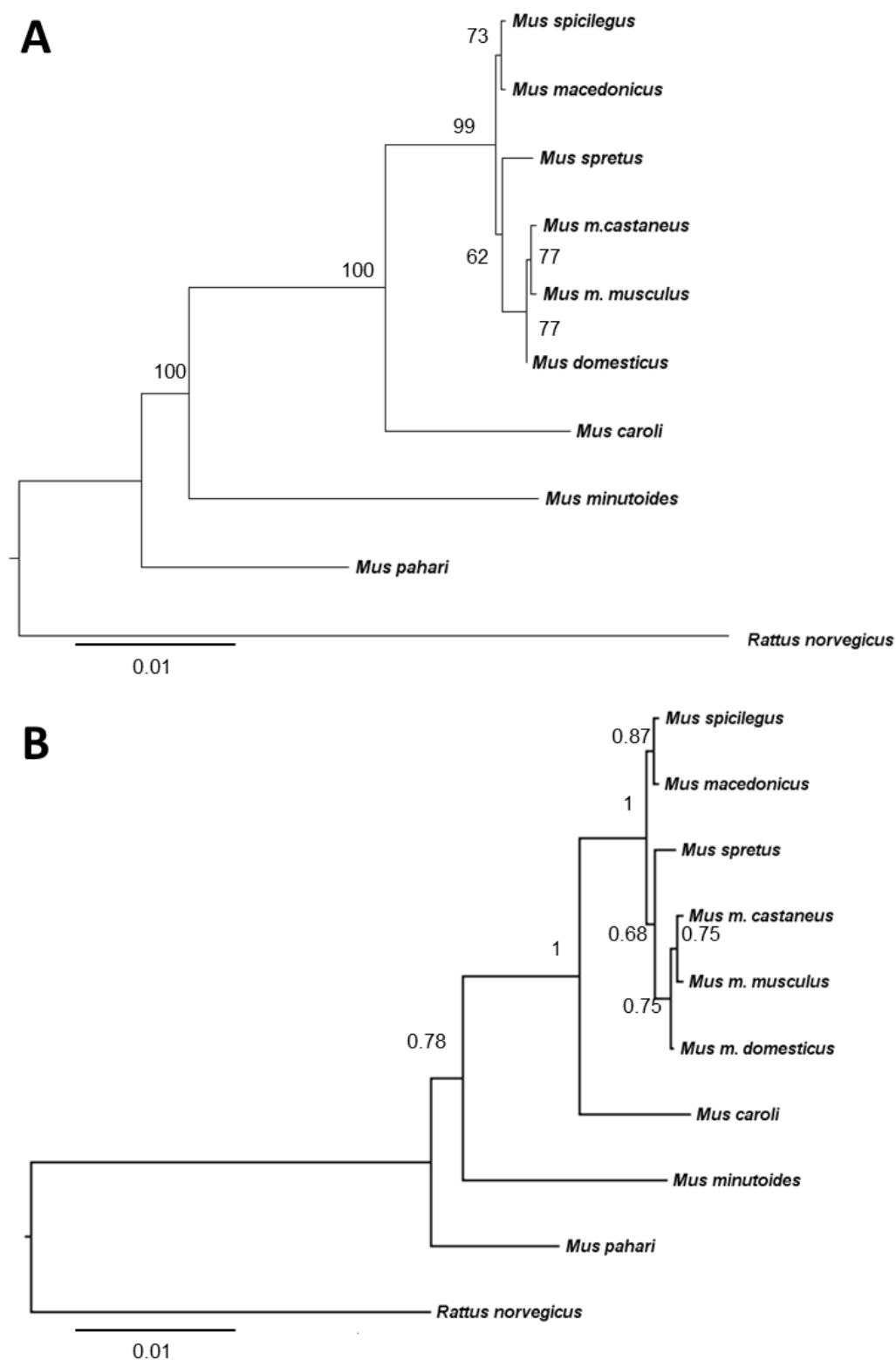


Figura 25. Árboles filogenéticos de *Izumo1* en roedores múridos. El soporte de cada rama tras un análisis de bootstrap no paramétrico de 100 réplicas así como la probabilidad Bayesiana posterior se representan en los nodos internos de las filogenias de Máxima Verosimilitud (**A**) e Inferencia Bayesiana (**B**) respectivamente. La longitud de las ramas indican las distancias genéticas, estimadas como el número de cambios sinónimos por codon. En el análisis Bayesiano, los parámetros se estimaron tras aplicar 100.000 generaciones y la frecuencia de muestreo fue de cada 100 generaciones. Las escalas se indican debajo de cada árbol.

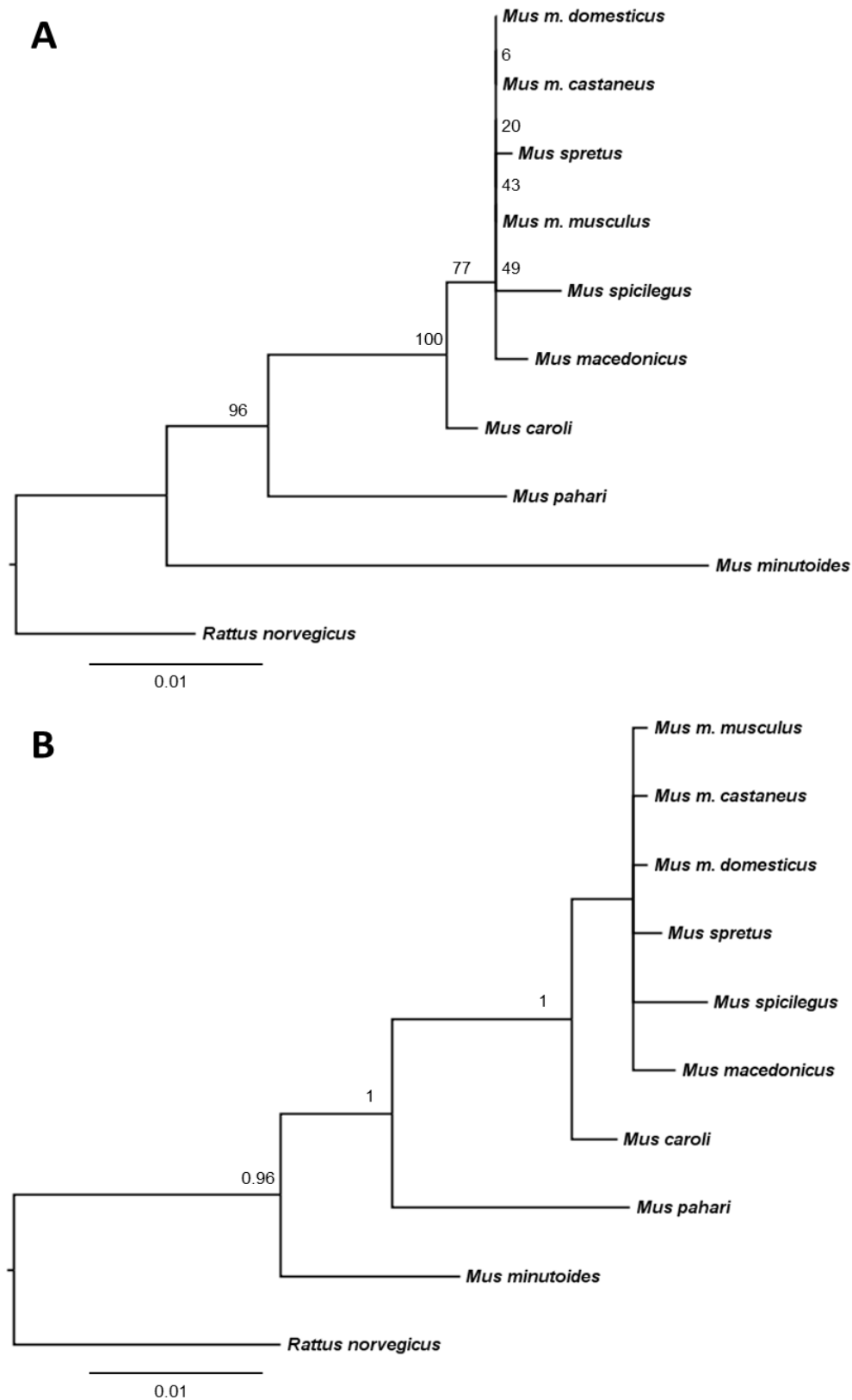


Figura 26. Árboles filogenéticos de *Cd9* en roedores múridos. El soporte de cada rama tras un análisis de bootstrap no paramétrico de 100 réplicas así como la probabilidad Bayesiana posterior se representan en los nodos internos de las filogenias de Máxima Verosimilitud (A) e Inferencia Bayesiana (B) respectivamente. La longitud de las ramas indican las distancias genéticas, estimadas como el número de cambios sinónimos por codon. En el análisis Bayesiano, los parámetros se estimaron tras aplicar 100.000 generaciones y la frecuencia de muestreo fue de cada 100 generaciones. Las escalas se indican debajo de cada árbol.

4.17. Análisis de selección positiva

Se evaluó si las proteínas Izumo1 y Cd9 se encuentran evolucionando bajo el efecto de la selección positiva mediante la aplicación de modelos capaces de calcular las tasas evolutivas (ω) para cada codon. Se detectó evidencia de selección positiva tanto en Izumo1 como en Cd9. En ambos casos, los LRTs fueron significativos en favor de los modelos de selección (M2 y M8) (Tabla XVII). Un total de 14 aminoácidos fueron identificados bajo selección positiva en el alineamiento de Izumo1 con un valor promedio de $\omega = 4,654$ (Fig. 22). En Cd9 sólo se detectaron dos residuos sujetos a selección positiva (Fig. 22), aunque el valor de ω estimado fue mayor que en Izumo ($\omega = 34,381$).

Tabla XVII. Análisis de selección positiva en Izumo1 y Cd9.

Proteína	N	L	Modelos comparados	LRT ^a	Estimas de parámetros	Aminoácidos bajo selección positiva ^b
Izumo1	9	376	M1a-M2a	47,511 **	p0=0,445, p1=0,384, p2=0,171, ω 0=0,105, ω 1=1, ω 2=4,612	42 (0,994**), 47 (0,999**), 49 (0,995**), 59 (0,957), 62 (0,997**), 70 (0,976*), 87 (0,975*), 150 (0,996**), 240 (0,970*), 343 (0,959*), 345 (0,970*)
			M7-M8	48,201 **	p0=0,830, p=0,830, q=0,139, p1=0,169, $\omega = 4,654$	42 (0,997**), 47 (1,000**), 48 (0,961), 49 (0,998**), 59 (0,978*), 62 (0,999**), 70 (0,988*), 87 (0,988*), 117 (0,962*), 127 (0,961*), 150 (0,998**), 240 (0,983*), 343 (0,979*), 345 (0,985*), 365 (0,972*)
			M8a-M8	47,737 **		
Cd9	9	225	M1a-M2a	15,045 **	p0=0,705, p1=0,295, ω 0 = 00, ω 1 = 1,000	153 (0,981*), 183 (0,971*)
			M7-M8	15,208 **	p0=0,986, p=32,817, q=99,000, p1=0,014, $\omega = 34,381$	153 (0,989*), 183 (0,984*)
			M8a-M8	15,033		

N: número de secuencias.

L: longitud de las secuencias (número de codones).

^aLRTs significativos con un nivel de significancia del 95% (*), 99% (**). Se muestran los parámetros estimados para el modelo que mejor se ajusta.

^bAminoácidos bajo selección positiva con la probabilidad posterior de un análisis Bayesiano Empírico (BEB). entre paréntesis. Los residuos con probabilidad > 0,95 (*) y > 0,99 (**) son mostrados.

4.18. Localización de los residuos bajo selección positiva en modelos estructurales

Tras examinar la disposición de los diferentes dominios de Izumo1, se observó que la mayor parte de los aminoácidos sujetos a selección positiva en esta proteína se concentran dentro del dominio Izumo (Fig. 27), una región extracelular implicada en la formación de homodímeros. De los cuatro residuos restantes sujetos a evolución adaptativa, tres se localizaron dentro de un dominio citosólico que regula la formación de multímeros (Fig. 27A). Con el fin de establecer la localización física de los residuos bajo selección positiva, se encontró un modelo estructural para Izumo1 basado en la estructura cristalográfica del dominio extracelular del receptor del factor de crecimiento ErbB1 (Bouyain *et al.* 2005) (Fig. 27B). Este modelo alinea con la región extracelular de Izumo1 (fragmento desde el residuo 32 al 294) con un 16% de identidad. En base a esta estructura, se observó que los aminoácidos bajo selección positiva encontrados a lo largo del dominio Izumo forman un cluster en la región expuesta al espacio extracelular.

Por otro lado, los dos codones que mostraron una señal significativa de selección positiva en Cd9 fueron localizados en el lazo grande extracelular (Fig. 27C). Esta región está implicada en la fusión del óvulo con el espermatozoide (Zhu *et al.* 2002). Como modelo estructural para el lazo grande extracelular de Cd9 se encontró la región correspondiente al lazo extracelular de la proteína Cd81 humana (Kitadokoro *et al.* 2001). La región coincidente mostró una identidad de secuencia del 24 % y abarcaba la región entre las posiciones 109 y 190 de Cd9. Dentro de la estructura tridimensional, los aminoácidos bajo selección positiva fueron encontrados en las regiones flanqueantes de los residuos SFQ, críticos para la fusión gamética (Zhu *et al.* 2002) (Fig. 27D).

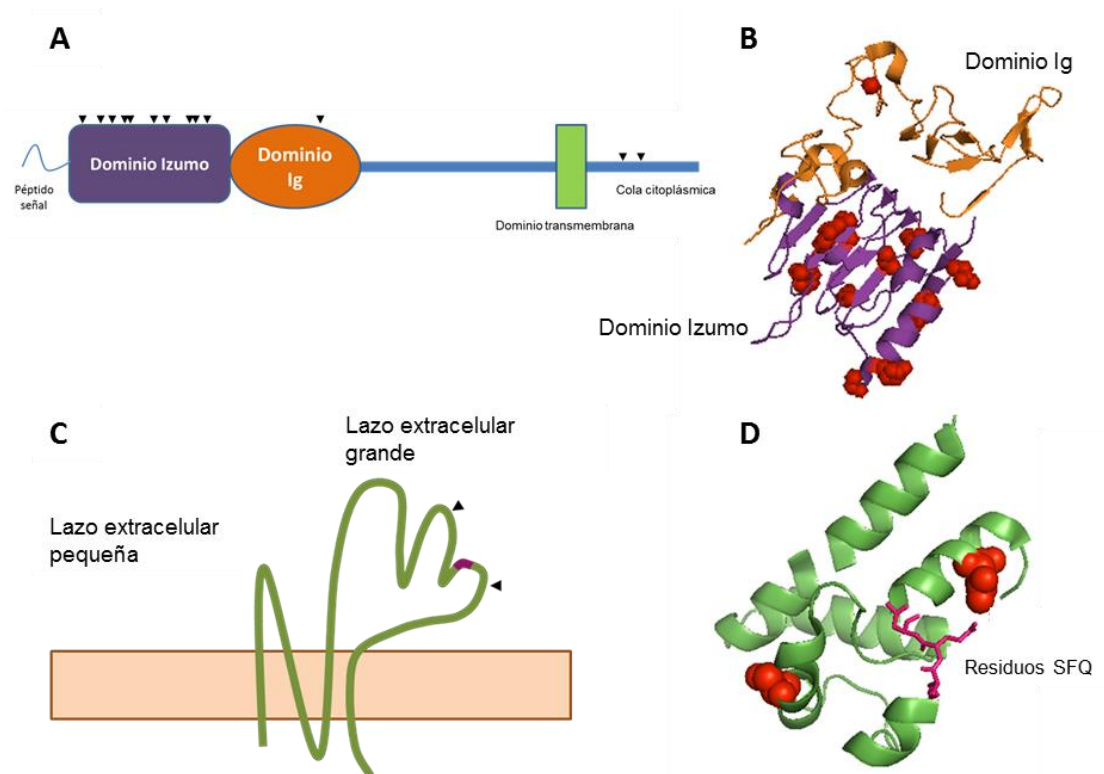


Figura 27. Estructura de Izumo1 y Cd9. **A.** Esquema de Izumo1 con sus dominios representados. **B.** Estructura tridimensional de la región extracelular de ErbB4 (referencia PDB: 2AHX), que sirve como modelo de la región extracelular de Izumo1. **C.** Esquema de Cd9 representando la distribución de los residuos bajo selección positiva. **D.** Modelo tridimensional del lazo extracelular grande de Cd9 basado en el dominio extracelular de Cd81 humana (ref. PDB: 1G8Q). Los residuos bajo selección positiva en los esquemas se muestran como flechas negras y en las estructuras tridimensionales como esferas rojas.

4.19. Coevolución entre *Izumo1* y *Cd9*

Se evaluó si las proteínas *Izumo1* y *Cd9* experimentan coevolución analizando si existe evidencia de correlación entre las tasas evolutivas de las proteínas durante la divergencia de las diferentes especies. El objetivo de estos análisis fue predecir si *Izumo1* y *Cd9* pueden tener una interacción directa basado en la idea de que un par de proteínas que interaccionan mostrarán historias evolutivas similares debido a los cambios compensatorios y a presiones selectivas mutuas (Pazos y Valencia 2008).

Se usó el valor de ω como medida de tasa evolutiva. Un valor de ω se estimó para cada rama del árbol filogenético incluyendo las 9 especies de móridos analizadas (Fig. 28) usando el modelo de rama libre de codeml. Las correlaciones entre los valores de ω de *Izumo1* y *Cd9* se llevaron a cabo mediante análisis de regresión. Para los análisis de coevolución, se probaron diferentes métodos de estimación de los valores de ω específicos de cada linaje. En una primera prueba se compararon los valores de ω estimados para cada rama terminal de la filogenia. Una desventaja de este método es que las ramas terminales frecuentemente carecen de sustituciones sinónimas porque tienen longitudes de rama muy cortas, lo que conlleva a la estimación de ω extremadamente altos o infinitos, siendo asignados a valores de 999 por el modelo. En el caso de *Izumo1*, la rama terminal de la especie *Mus*

m. castaneus mostró un ω infinito. Con tal de evitar el efecto disruptivo de estos valores arbitrarios sobre los análisis de regresión, estos se excluyeron del análisis. Los resultados de la regresión lineal usando solo los valores definidos de ω fueron no significativos ($F = 2,739$, $p = 0,149$). En un intento por incluir todos los linajes en los análisis, se combinaron aquellas ramas terminales que muestran valores ω extremos o infinitos con la rama más adyacente con el fin de obtener una ω estimada desde el ancestro más reciente (Dorus *et al.* 2004). Con este método tampoco se encontró evidencia entre las tasas evolutivas de las dos proteínas ($F = 0,3249$, $p = 0,586$).

Teniendo en cuenta los inconvenientes que presenta el uso de valores de ω calculados para las ramas terminales en los análisis evolutivos, se aplicó un método basado en calcular los valores de ω desde la raíz del árbol hasta cada rama terminal. Esta metodología presenta algunas ventajas para los análisis de coevolución. Por ejemplo, hace que se considere la historia evolutiva completa de los *loci* analizados, mejorando la medida de simetría entre sus árboles filogenéticos, que es interpretado como una señal de evolución coordinada entre un par de proteínas (Pazos y Valencia 2001). Por otra parte, la estima de tasas evolutivas desde el ancestro común transforma las ramas para que tengan la misma longitud y, por tanto, hace que los análisis no se encuentren sometidos a efectos temporales (Wolf *et al.* 2009).

La estima de ω específicos para cada linaje por este método resultó en valores próximos a 1 para todas las especies (Fig. 28). Estos datos vienen a confirmar que Izumo1 es una proteína con una tasa de evolución elevada. En el caso de Cd9, se estimaron valores de ω más bajos (Fig. 28), apoyando la evidencia de que esta proteína diverge más lentamente que Izumo1.



Figura 28. Tasas evolutivas de Izumo1 y Cd9. *Rattus norvegicus* se utilizó como grupo externo para establecer la raíz de las filogenias. La topología de los árboles representados está basada en filogenias previamente establecidas de roedores (Suzuki *et al.* 2005; Fabre *et al.* 2012).

Se obtuvo una correlación significativa positiva entre las estimas de ω Izumo1 y Cd9 calculadas desde la raíz ($F = 29,19$, $p = 0,001$) (Fig. 29), lo que muestra una señal de coevolución entre este par de proteínas. Dado que el dominio Izumo extracelular de Izumo1 y el lazo extracelular grande de Cd9 fueron las regiones mostrando la señal de selección positiva más intensa, se evaluó si la divergencia de estos dominios es lo que realmente dirige la coevolución de estas dos proteínas. Para ello, se estimaron los valores de ω específicos de cada linaje también para los alineamientos del dominio Izumo y del lazo extracelular grande de Cd9. Se obtuvo una correlación positiva cuando las tasas evolutivas del dominio Izumo se compararon con la secuencia completa de Cd9 ($F = 18,272$, $p = 0,0036$). Por otro lado, las estimas de ω del lazo extracelular grande de Cd9 mostraron una correlación positiva cuando se compararon con la secuencia completa de Izumo1 ($F = 69,569$, $p = 0,00007$). Cuando las tasas evolutivas de estas dos regiones fueron comparadas, se obtuvo una correlación positiva con un nivel de significancia más elevado ($F = 95,512$, $p = 0,00002$) (Fig. 30). Con el fin de confirmar que la señal de coevolución es originada principalmente por estas dos regiones, también se analizó la coevolución de otros dominios de Izumo1 (ej: el dominio Ig o la cola citoplásmica) y Cd9 (ej: el lazo extracelular pequeño), pero no se obtuvieron relaciones significativas en ningún caso (Fig. 30).

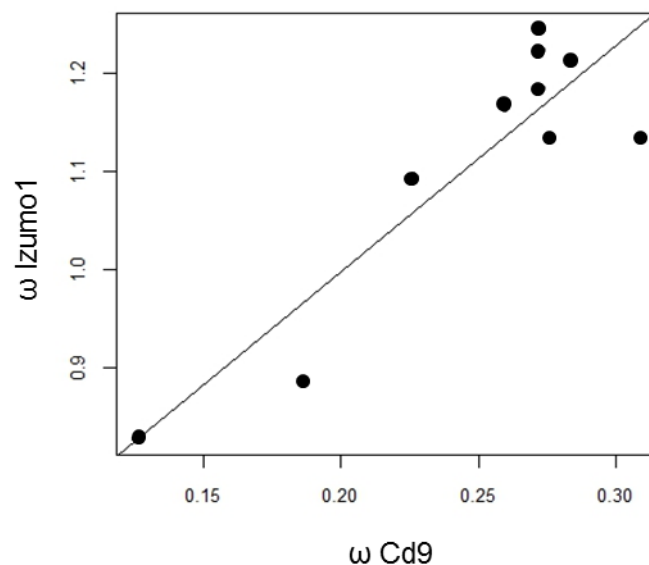


Figura 29. Coevolución entre Izumo1 y Cd9. Las tasas evolutivas (ω) se estimaron para 9 especies de roedores móridos. La diferencia entre las escalas revela la mayor tasa evolutiva de Izumo1.

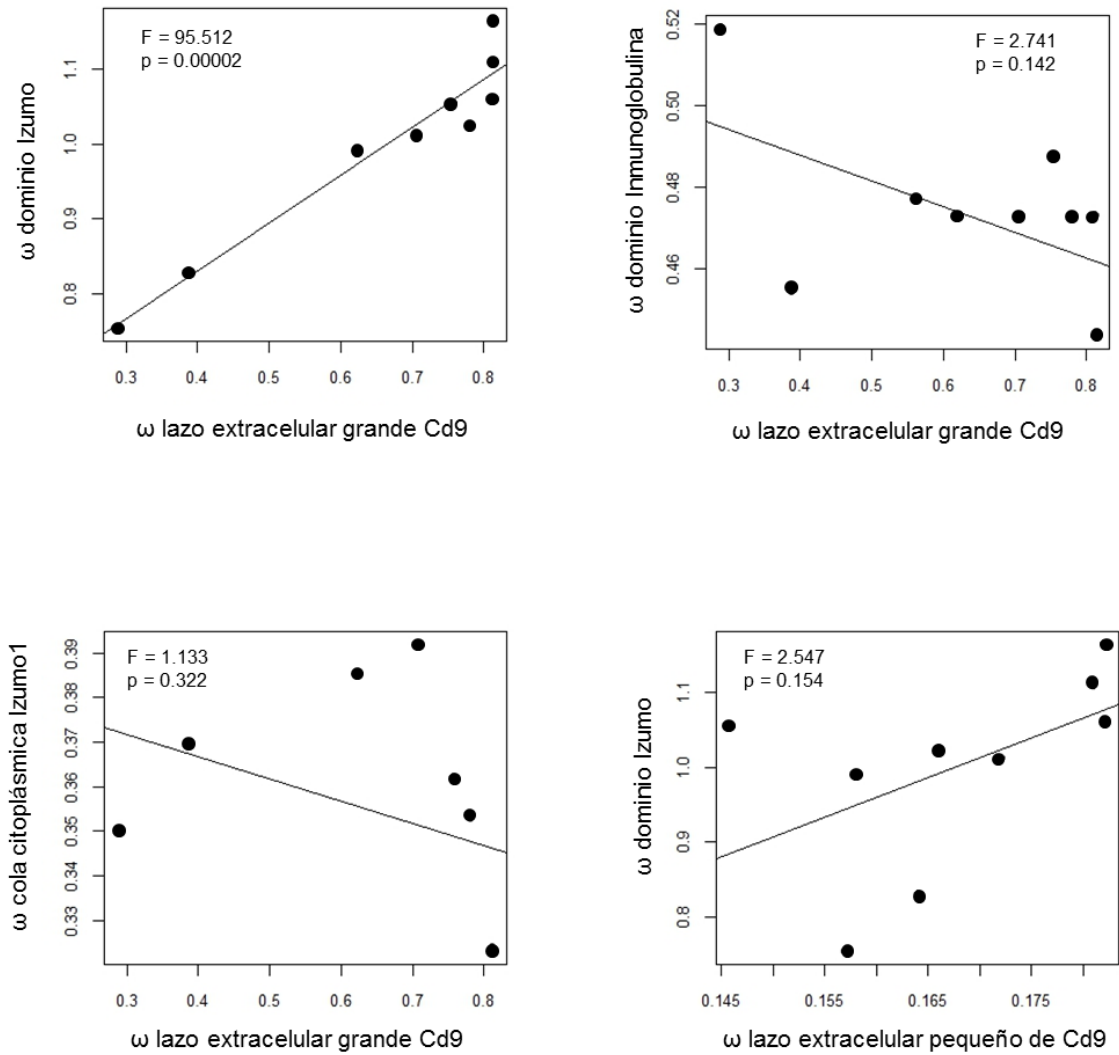


Figura 30. Análisis de regresión entre los dominios de Izumo1 y Cd9. Se muestran los resultados estadísticos en cada gráfica.

Con el fin de evaluar la robustez de los resultados de los análisis de coevolución entre Izumo1 y Cd9, se aplicaron dos tipos de enfoque. Por una parte, se evaluó la influencia que puede tener el utilizar una filogenia de especies reconstruida. Para ello, se realizaron análisis de coevolución aplicando filogenias con diversas topologías. En primer lugar, se analizó el efecto de una filogenia reconstruida a partir de un alineamiento conjunto de Izumo1 y Cd9 por los métodos de Máxima Verosimilitud e Inferencia Bayesiana. Estas metodologías produjeron árboles con diferentes topologías entre sí y respecto a la filogenia reconstruida. No obstante, las correlaciones entre las ω de Izumo1 y Cd9 siguieron siendo significativas (Fig. 31A, B). En segundo lugar, se exploró el efecto de una topología que incluía la politomía frecuente entre las tres subespecies de *M. musculus*. Utilizando esta topología, los resultados de coevolución entre Izumo1 y Cd9 se mantuvieron como significativos. (Fig. 31C). Por otra parte, se analizó la posibilidad de que la coevolución observada

entre Izumo1 y Cd9 sea consecuencia de un falso positivo. Para ello, se llevaron a cabo análisis de coevolución con diversas proteínas que no tienen funciones relacionadas con las de Izumo1 y Cd9. Las proteínas elegidas fueron Catsper1, PKDREJ y Prm2. Los análisis de coevolución de estos genes no mostraron correlaciones significativas ni con Izumo1 ni con Cd9 (Tabla XVIII). También se analizó la coevolución con una proteína implicada en la fecundación de gametos y que se piensa que interacciona con Izumo1: Spesp1 (Fujihara *et al.* 2010). Sin embargo, esta proteína no mostró señales de coevolución con Izumo1 o Cd9 (Tabla XVIII).

Tabla XVIII. Análisis de coevolución de Izumo1 y Cd9 con proteínas no relacionadas.

Proteína Control	Proteína de análisis	Pendiente	R²	F	p
Catsper1 ^a	Izumo1	-1,181	0,245	2,265	0,176
	Cd9	-0,536	0,333	3,489	0,104
Pkdrej ^b	Izumo1	1,145	0,038	0,236	0,644
	Cd9	0,554	0,078	0,508	0,503
Prm2 ^c	Izumo1	-1,448	0,445	6,612	0,042
	Cd9	-0,451	0,356	4,866	0,069
Spsp1 ^d	Izumo1	-0,179	0,006	0,04	0,848
	Cd9	-0,355	0,147	1,202	0,309

^aSecuencia del primer exón del gen *Catsper1* (ver Bloque II).

^bSecuencia del dominio extracelular REJ de PKDREJ (ver Bloque III).

^cLa secuencia codificante de Prm2 (~445 pb) se amplificó por PCR estándar del ADN genómico..

^dLa secuencia codificante de Spesp1 (~1.197 pb) se amplificó por RT-PCR estándar de ADN complementario

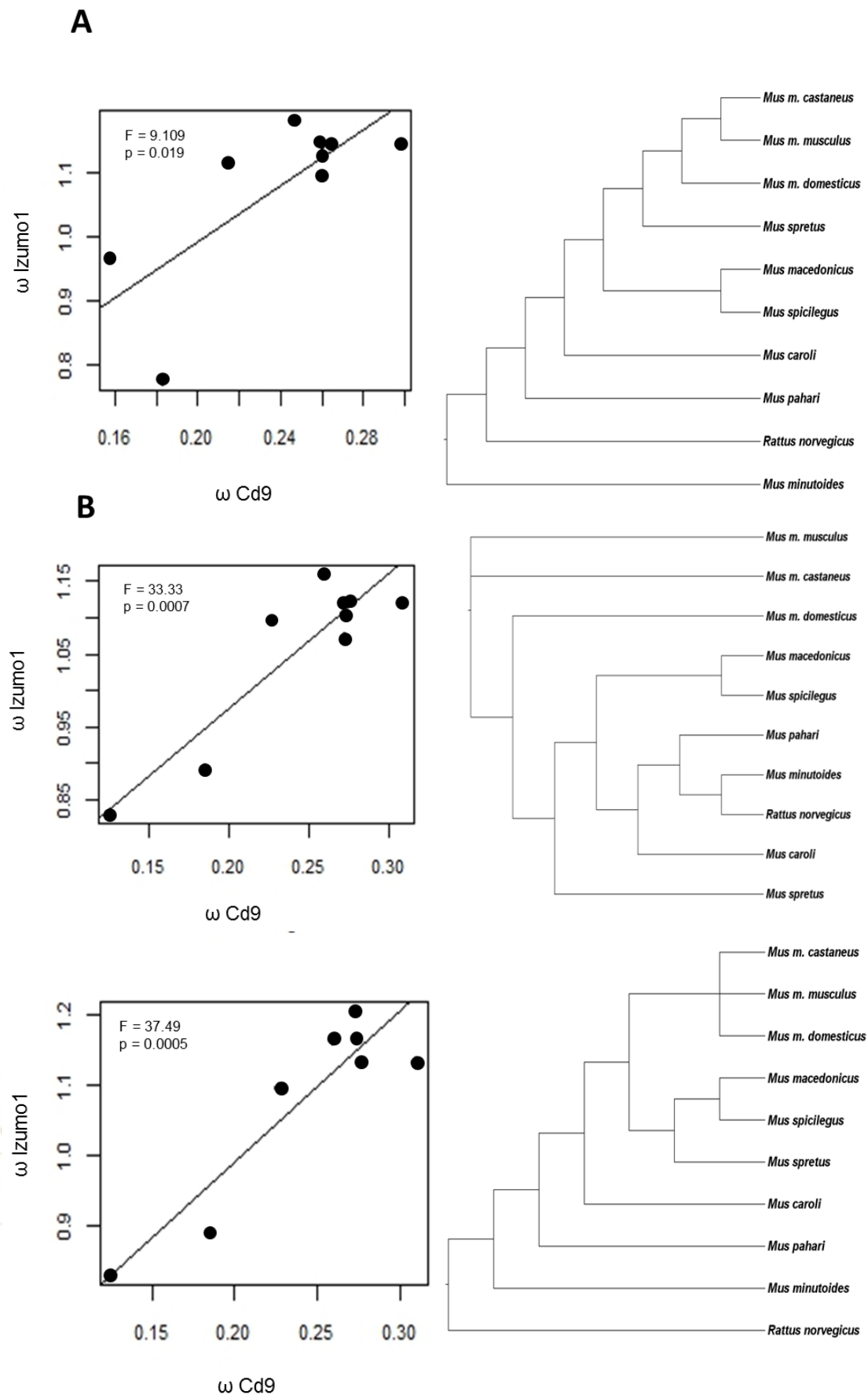


Figura 31. Efecto de la filogenia sobre los análisis de coevolución Izumo1-Cd9. Se evaluaron filogenias inferidas de un alineamiento de Izumo1 + Cd9 por Máxima Versomilitud (**A**) e Inferencia Bayesiana (**B**) y una filogenia con la politomia de las subespecies de *Mus* (**C**). Se muestran los resultados estadísticos en cada gráfica.

4.20. Comparación de sustituciones de *Izumo1* y *Cd9* en los linajes

La divergencia relativa entre un par de proteínas que interaccionan puede dar información sobre las fuerzas selectivas que están actuando sobre las especies en el presente (Civetta 2003b). Por ello, se estimó el número de sustituciones no sinónimas en las filogenias de *Izumo1* y *Cd9* y estos valores se compararon nodo por nodo (Fig.32). Se detectó un mayor número de cambios para *Izumo1* que para *Cd9* en la mayoría de los nodos. No obstante, las ramas terminales de *Mus macedonicus* y *Mus spicilegus* mostraron cambios en la secuencia de *Cd9* pero ninguno en la secuencia de *Izumo1*. Esto indica que las tasas de evolución relativas entre *Izumo1* y *Cd9* varían entre los linajes.

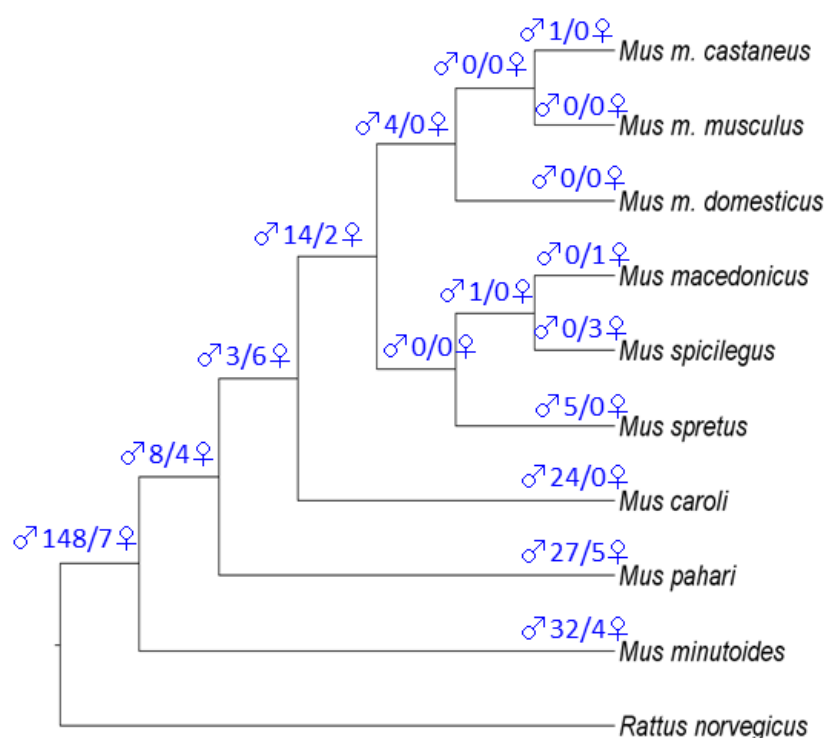


Figura 32. Número de cambios no sinónimos en *Izumo1* (símbolo masculino) y *Cd9* (símbolo femenino) a lo largo de una filogenia de 9 especies de muridos.

4.21. Asociación entre la evolución de *Izumo1* y *Cd9* y los niveles de competición espermática

Con el fin de analizar si la selección sexual post-cópula dirige la evolución de *Izumo1* y *Cd9*, se usó la masa testicular relativa a la masa corporal como índice de competición espermática. No se observaron relaciones significativas de esta variable con las tasas evolutivas de *Izumo1* ($F = 1,834$, $p = 0,229$) o de *Cd9* ($F = 0,27$, $p = 0,619$).

BLOQUE V

Análisis de proteómica cuantitativa revelan diferencias en la abundancia de proteínas del espermatozoide entre especies de ratón con diferente nivel de competición espermática.

Los espermatozoides presentan una baja actividad transcripcional y traduccional, por lo que los procesos que experimentan estas células para adquirir su funcionalidad y actividad desde su diferenciación en los testículos van a producirse mayormente mediante modificaciones post-traduccionales. Por lo tanto, el espermatozoide es un modelo importante para estudios de proteómica. Además, el espermatozoide es un factor clave desde el punto de vista evolutivo, ya que estas células son el foco de acción de fuerzas de selección sexual, como la competición espermática o el conflicto sexual. Estas presiones selectivas pueden causar la rápida evolución de *loci* reproductivos que da como resultado la adquisición o la adaptación de rasgos reproductivos que incrementan la eficacia de fecundación del espermatozoide.

En este estudio, se utilizaron dos técnicas de proteómica cuantitativa (2D-DIGE y LC-MS/MS con conteo de espectros no marcados) para analizar y comparar el proteoma del espermatozoide entre tres especies de roedores cercanas filogenéticamente pero que presentan diferentes niveles de competición espermática.

Los resultados de estos análisis revelaron un conjunto de proteínas que muestran variaciones en el nivel de abundancia entre estas especies. Varias proteínas mostraron mayor abundancia en las dos especies con altos niveles de competición espermática (*Mus spretus* y *Mus spicilegus*) en comparación con la especie de baja competición espermática (*M. musculus*). La mayoría de estas proteínas están implicadas en procesos importantes para el éxito de la fecundación, tales como la motilidad espermática, el metabolismo del espermatozoide, la reacción acrosómica o la interacción espermatozoide-óvulo.

4.22. Composición del proteoma del espermatozoide de ratón

Las proteínas del espermatozoide fueron digeridas y analizadas por métodos de espectrometría de masas (LC-MS/MS) de alta resolución (*ver* Material y Métodos). El número combinado de espectros MS/MS identificados como péptidos entre los tres replicados de cada especie fue de 15.018 para *M. musculus*, 16.244 para *Mus spretus* y 17.322 para *Mus spicilegus*. A partir de los datos crudos de los análisis de MS/MS, se identificaron un total de 1.989 proteínas para *M. musculus*, 1.944 para *M. spretus* y 2.073 para *M. spicilegus*, aunque finalmente se validaron 1.156 (*M. musculus*), 1.215 (*M. spretus*) y 1.205 (*M. spicilegus*) proteínas por los algoritmos de PeptideProphet™ y ProteinProphet™, tras aplicar una tasa de falso descubrimiento del 1% y eliminar las proteínas identificadas por un único péptido (Tablas suplementarias 7-9).

Cuando los datos de los análisis de MS/MS se compararon entre las réplicas de cada especie, se identificó un número variable de proteínas entre estas, aunque el número de proteínas solapantes fue constante entre las especies. Un total de 858 (74,2%) proteínas se identificaron en al menos dos réplicas de *M. musculus*, 973 (80%) en *M. spretus*, y 976 (81%) en *M. spicilegus* (Fig. 33, Tablas suplementarias 7-9). Con el fin de considerar los cambios que se han podido producir desde la separación de las especies, los análisis de identificación de proteínas también se realizaron teniendo en cuenta mutaciones puntuales en las secuencias de los péptidos detectados. De este modo, se identificaron 2 proteínas adicionales en *M. musculus*, 15 en *M. spretus* y 7 en *M. spicilegus* (Tablas suplementarias 7-9). A nivel interespecífico, se observó que un total de 690 (58,1%) proteínas se identificaron con consistencia (es decir, en dos o más réplicas por especie) en las 3 especies. De entre estas, 400 (37,7%) proteínas se detectaron en las 3 réplicas de las 3 especies.

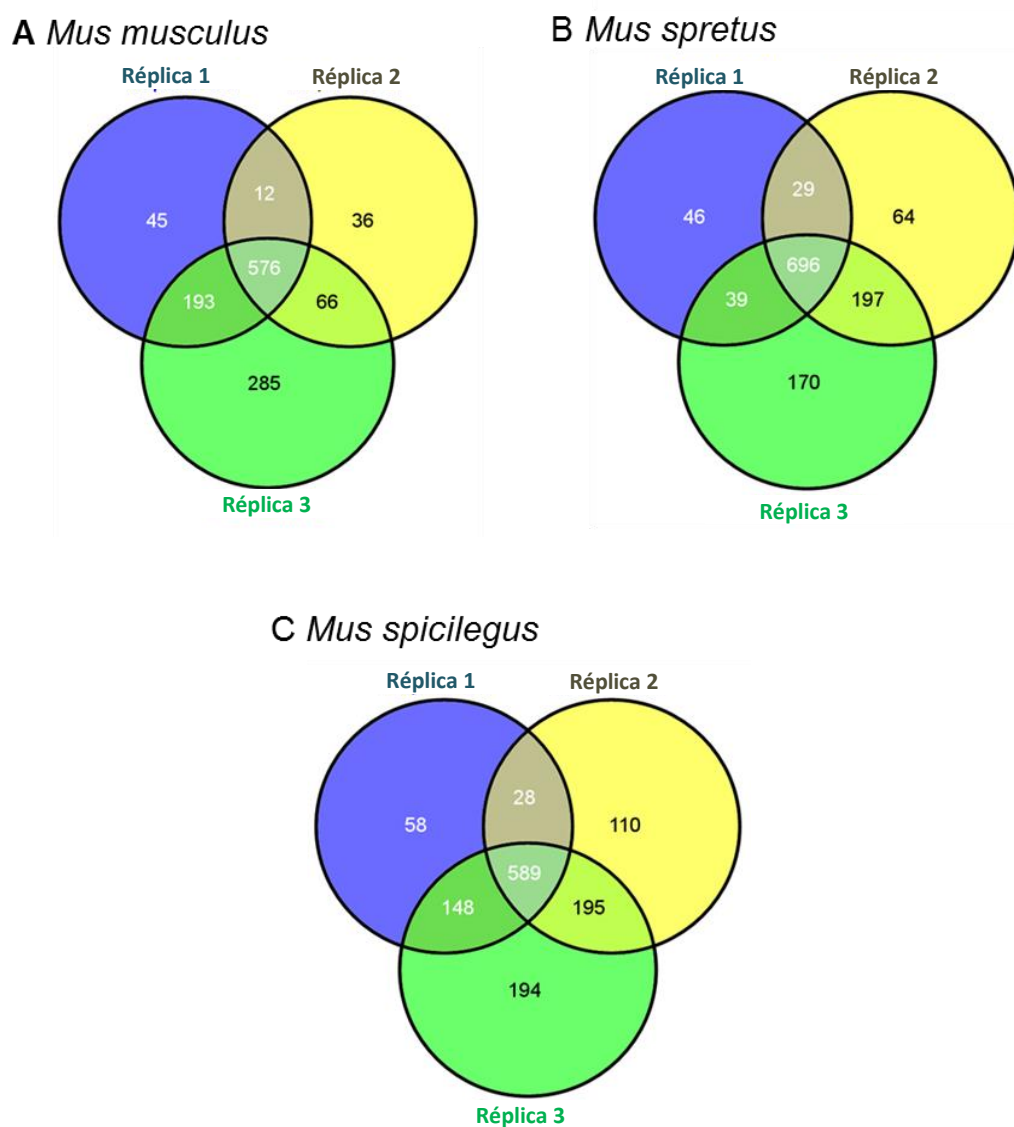


Figura 33. Representación en diagramas Venn de las proteínas identificadas mediante espectrometría de masas en las 3 réplicas de cada especie. Los diagramas fueron diseñados con la herramienta Venny (Oliveros 2007).

4.23. Anotación funcional del proteoma del espermatozoide de ratón

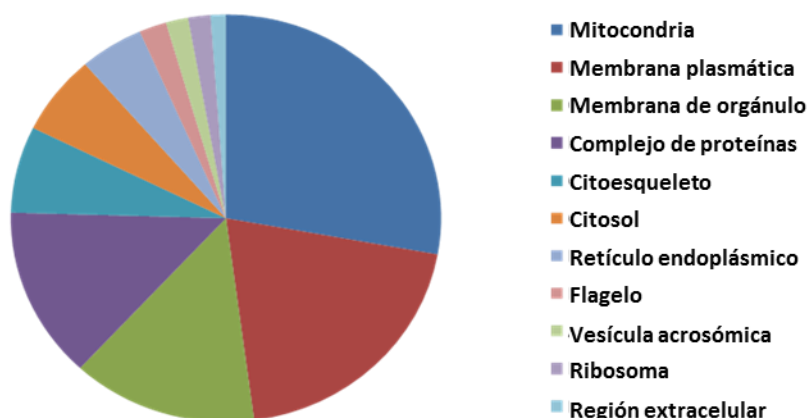
Las proteínas identificadas por los análisis de MS se clasificaron en categorías funcionales establecidas por *Gene Ontology*. Los proteomas de las tres especies de ratón revelaron una distribución casi idéntica de proteínas entre los diferentes grupos, dando evidencia de la alta reproducibilidad de los métodos basados en MS. Por esta razón, y con el fin de simplificar la descripción de los resultados, en esta sección solo se detallan los resultados para la especie *M. musculus*.

Un total de 843 proteínas fueron anotadas para componentes subcelulares, revelando un enriquecimiento de proteínas mitocondriales (27,8%). El método de LC-MS/MS empleado fue altamente eficaz en el aislamiento de proteínas de la membrana, identificando un total de 263 (20,1%) proteínas, 207 de las cuales son proteínas integrales de membrana (Fig. 34). Respecto a los procesos biológicos, más de la mitad de las proteínas fueron asignadas a procesos metabólicos (463 de un total de 837, 53,6%). También se detectó un número sustancial de proteínas participando en el transporte intracelular (17,2%) (Fig. 34). La clasificación en términos de la función molecular reveló que las proteínas catalíticas constituyen la categoría predominante (381 de un total de 786, 38,5%) (Fig. 34). Un amplio número de proteínas se asignaron a funciones de unión (20,3%), la mayoría probablemente implicadas en unión a ácidos nucleicos (Wasbrough *et al.* 2010). Las proteínas estructurales tuvieron una representación del 9,7%.

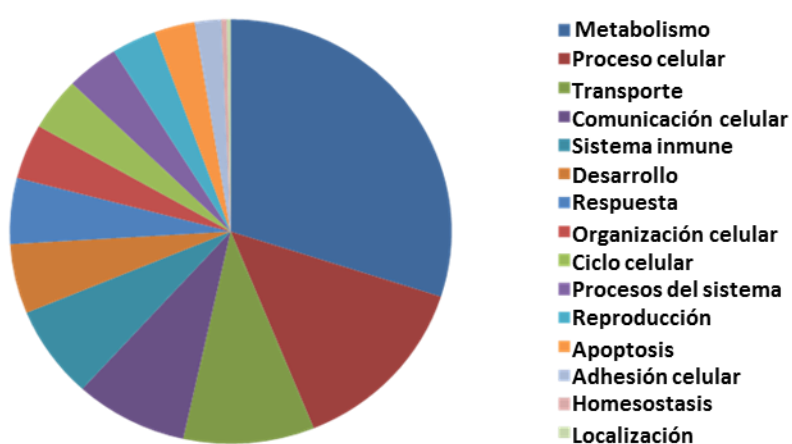
Usando el proteoma completa de ratón como *background*, se identificaron un total de 218 clusters significativos de proteínas para *M. musculus*, 202 para *M. spretus* y 220 para *M. spicilegus*. Los 25 clusters mejor representados se muestran en la tabla XIX. La mayoría de estos clusters se encuentran asociados con procesos biosintéticos y metabólicos. Las proteínas implicadas en el plegamiento de proteínas y el reconocimiento gamético también formaron grupos enriquecidos. Estos resultados indican la importancia de un metabolismo activo, un sistema de transporte eficiente y una regulación estricta del plegamiento de proteínas en el espermatozoide.

Un total de 107 proteínas identificadas en nuestro estudio mostró un fenotipo mutante asociado con anomalías en la reproducción masculina (Tabla suplementaria 10). Se identificaron fenotipos reproductivos muy variados como anomalías en la meiosis o la espermatogénesis, morfología anormal de las células germinales o del espermatozoide, defectos en la motilidad del espermatozoide, fallos en la capacitación o la reacción acrosómica, y una reducción o pérdida de la fertilidad.

Componente subcelular



Proceso biológico



Función molecular

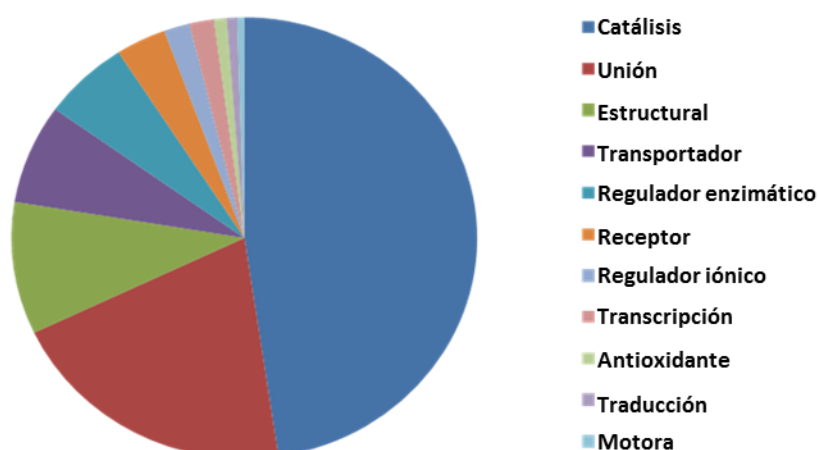


Figura 34. Diagramas representando la clasificación de las proteínas identificadas en el proteoma del espermatozoide de *Mus musculus*.

Tabla XIX. Lista de los 25 procesos biológicos más enriquecidos en el proteoma del espermatozoide de ratón.

Término de anotación GO	<i>M. musculus</i>		<i>Mus spretus</i>		<i>Mus spicilegus</i>	
	<i>p</i>	Número de proteínas	<i>p</i>	Número de proteínas	<i>p</i>	Número de proteínas
Proceso metabólico del organismo	5,11E-66	317	5,07E-71	333	8,86E-71	345
Proceso metabólico de pequeñas moléculas	1,13E-48	217	8,99E-49	222	2,75E-49	231
Proceso de oxidación-reducción	1,70E-47	153	5,83E-51	161	1,85E-54	171
Proceso metabólico	2,15E-32	555	4,84E-33	574	3,32E-32	600
Proceso metabólico del ácidos carboxílicos	1,44E-28	107	7,79E-28	108	1,29E-28	113
Proceso metabólico de ácidos orgánicos	3,73E-27	110	8,10E-26	110	1,89E-26	115
Proceso metabólico de oxiácidos	1,15E-26	108	6,38E-26	109	1,38E-26	114
Proceso catabólico	1,14E-25	159	6,05E-31	174	4,25E-31	181
Proceso metabólico de compuestos organonitrogenados	3,54E-25	145	5,66E-25	148	5,84E-27	158
Proceso catabólico de sustancias orgánicas	9,00E-24	144	8,68E-29	158	2,43E-27	161
Proceso metabólico de organofosfatos	2,71E-21	111	1,55E-20	112	2,25E-21	118
Proceso metabólico de moléculas conteniendo nucleobases	4,91E-21	92	3,96E-22	96	1,86E-22	100
Proceso metabólico de nucleótidos	1,39E-20	87	3,30E-21	90	3,16E-22	95
Plegamiento de proteínas	2,02E-20	38	4,83E-18	36	5,32E-19	38
Proceso metabólico de nucleósidos trifosfato	3,41E-20	87	8,35E-21	90	8,45E-22	95
Proceso biosintético de nucleósidos trifosfato	9,09E-20	26	2,25E-19	26	1,48E-17	25
Ciclo de ácidos carboxílicos	1,51E-19	19	3,00E-19	19	2,03E-20	20
Proceso catabólico celular	1,33E-18	126	1,04E-20	134	2,53E-22	143
Proceso metabólico de compuestos glicosilados	7,25E-18	70	4,97E-17	70	1,83E-18	75
Proceso metabólico de sustancias orgánicas	2,21E-17	467	2,01E-17	482	1,63E-16	503
Proceso catabólico de glucosa	8,78E-17	22	2,99E-20	25	1,24E-19	25
Reconocimiento espermatozoide-óvulo	9,12E-17	19	7,00E-18	20	5,32E-16	19
Proceso biosintético de ribonucleósidos trifosfato	1,56E-16	21	3,24E-16	21	1,06E-15	21
Proceso metabólico de cofactores	1,57E-16	46	2,40E-17	48	1,85E-18	51
Proceso biosintético de nucleósidos trifosfato de purina	2,78E-16	21	5,80E-16	21	1,89E-15	21

4.24. Detección de proteínas con presencia exclusiva en las especies

Se trató de identificar aquellas proteínas con espectros detectados en solo una de las especies. Con el fin de garantizar el manejo de datos robustos, se seleccionaron aquellas proteínas para las que se encontraron espectros de MS/MS en al menos dos de las tres réplicas de una especie pero ningún espectro en ninguna réplica del resto de las especies. Aplicando este criterio, se encontraron 26 proteínas con espectros exclusivamente detectados en *M. musculus*, 23 en *M. spretus*, y 30 en *M. spicilegus* (Tabla suplementaria 11).

4.25. Análisis cuantitativos de APEX

La cuantificación de proteínas analizadas por LC-MS/MS fue llevada a cabo usando la herramienta de proteómica cuantitativa APEX. Los datos de abundancia se calcularon para cada proteína computando el número combinado de espectros entre las tres réplicas de cada especie. Las tres especies mostraron una distribución idéntica de los datos cuantitativos de APEX (Figura suplementaria 8). Las proteínas con mayor abundancia fueron enzimas implicadas en metabolismo y producción de energía, mediadores de la cadena transportadora de electrones, y un grupo de proteínas con función estructural en el flagelo del espermatozoide (Tabla XX).

Tabla XX. Lista de las 25 proteínas más abundantes en el espermatozoide de ratón.

Proteína	Nombre de la proteína	APEX score			Número de espectros		
		<i>M. musculus</i>	<i>Mus spretus</i>	<i>Mus spicilegus</i>	<i>M. musculus</i>	<i>Mus spretus</i>	<i>Mus spicilegus</i>
SPACA1	sperm acrosome membrane-associated protein 1 precursor	0,015771523	0,011496927	0,015343843	294	278	330
LDHC	L-lactate dehydrogenase C chain	0,010370016	0,011579181	0,012658673	716	966	928
COX2	cytochrome c oxidase subunit II	0,00974733	0,007415808	0,008015447	98	81	76
Gm6123	pyruvate dehydrogenase E1 component subunit beta,	0,008956886	0,007517306	0,008671776	437	473	445
RAB2A	ras-related protein Rab-2A	0,008902561	0,008857603	0,010802675	292	352	375
COX4I1	cytochrome c oxidase subunit 4 isoform 1, mitochondrial precursor	0,008421194	0,005429986	0,005993724	215	154	165
ATP5B	ATP synthase subunit beta, mitochondrial precursor	0,00822562	0,009936516	0,010215866	668	948	864
Gm128	protein MENT precursor	0,008122374	0,009328123	0,009205995	220	280	250
SPESP1	sperm equatorial segment protein 1 precursor	0,008001396	0,009194719	0,007310148	315	398	302
LOC100046138	succinyl-CoA:3-ketoacid coenzyme A transferase 2B,	0,007713328	0,009599666	0,008633791	762	1097	888
PDHA2	pyruvate dehydrogenase E1, testis-specific form,	0,007585223	0,007200256	0,006577202	377	435	354
GAPDHS	glyceraldehyde-3-phosphate dehydrogenase, testis-specific	0,006845997	0,007178342	0,005258094	415	513	336
PKM	pyruvate kinase, muscle isoform M1	0,006395322	0,006761807	0,005653915	627	800	581
UQCRC2	cytochrome b-c1 complex subunit 2, mitochondrial precursor	0,006359046	0,007346425	0,005865026	303	402	293
ATP5A1	ATP synthase subunit alpha, mitochondrial precursor	0,006271242	0,007283678	0,006031229	539	742	602
LDHAL6B	L-lactate dehydrogenase A-like 6B	0,006196426	0,005257715	0,006514982	391	370	424
SUCLG1	succinyl-CoA ligase [ADP/GDP-forming] subunit alpha,	0,006051398	0,00943364	0,00893407	157	276	251
AKAP4	A-kinase anchor protein 4 isoform b	0,005859698	0,005202227	0,002431018	831	788	368
DLAT	dihydrolipoyllysine-residue acetyltransferase	0,005802588	0,005934628	0,004652064	425	547	384
Gm16399	cytochrome c oxidase subunit 6C	0,005556151	0,004448522	0,005040575	57	45	56
SLC2A3	solute carrier family 2, facilitated glucose transporter member 3	0,00542572	0,004831955	0,003357845	137	148	93
ACO2	aconitate hydratase, mitochondrial precursor	0,005409188	0,004188298	0,004379714	495	436	417
ACRV1	acrosomal protein SP-10 precursor	0,005284914	0,006702331	0,005655238	53	76	63
Gm3982	mitochondrial pyruvate carrier 2	0,005065002	0,004985642	0,004475488	61	68	58
Gm5506	alpha-enolase	0,005055182	0,004177317	0,004099178	318	320	280

4.26. Análisis comparativos de los datos cuantitativos

Los datos cuantitativos estimados con APEX fueron comparados entre las especies de roedores mediante un análisis Z-score, usando un punto de corte para la tasa de falsos descubrimientos del 5%. Se compararon un total de 1.057 proteínas entre *M. musculus* y *M. spretus* y se detectaron diferencias significativas entre los valores de APEX para 50 proteínas ($p < 0,0005$ después de aplicar la corrección de Bonferroni). Dentro de estas, 30 proteínas mostraron mayor abundancia en *M. musculus* y 20 en *M. spretus* (Fig. 35, Tabla suplementaria 12). En la comparación *M. musculus* vs *M. spicilegus*, se examinaron un total de 1.085 proteínas, revelando 26 proteínas con una abundancia mayor en *M. musculus* contra 21 proteínas con mayor abundancia en *M. spicilegus* (Fig. 35, Tabla suplementaria 13). En el análisis de *M. spretus* vs *M. spicilegus* se compararon 1.139 proteínas, con 20 de ellas mostrando mayor abundancia en *M. spretus* y 13 en *M. spicilegus* (Fig. 35, Tabla suplementaria 14).

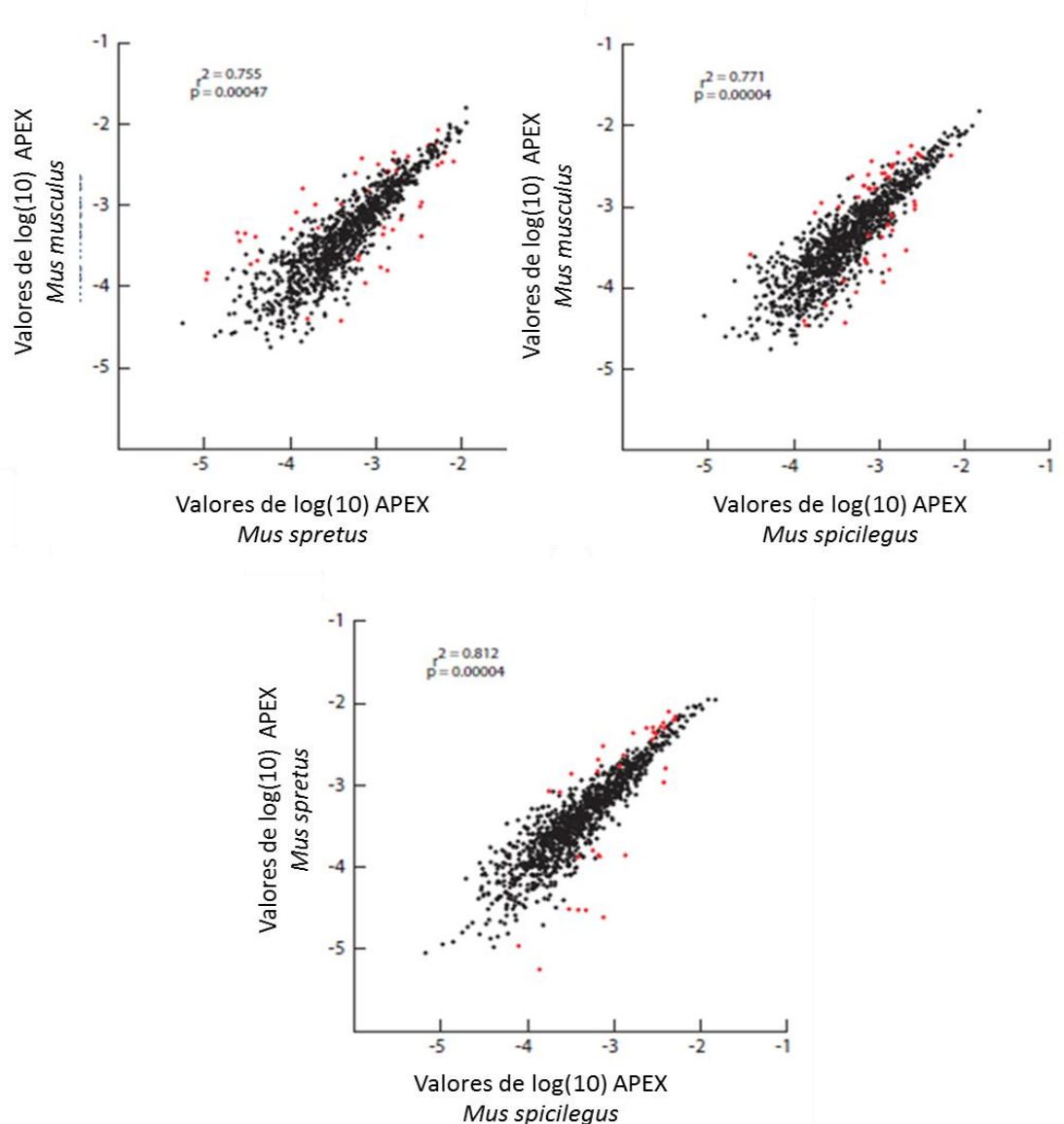


Figura 35. Gráficas mostrando la correlación de los valores de abundancia de APEX entre las especies de roedores analizadas. Los datos se transformaron a escala logarítmica. Los puntos rojos representan las proteínas con diferencias significativas.

4.27. *Análisis de péptidos únicos*

Dado que los péptidos detectados por MS se identificaron contra una base de datos correspondiente a la especie *M. musculus*, los análisis para *M. spretus* y *M. spicilegus* se llevaron a cabo por identificación entre especies. Por lo tanto, se comprobó que las proteínas que mostraron un número de espectros significativamente diferente no eran el resultado de falsos positivos debidos a la detección exclusiva de péptidos con secuencias diferentes entre las especies. Para ello, se identificaron todos aquellos péptidos que fueron detectados exclusivamente entre las especies comparadas. En la comparación entre *M. musculus* y *M. spretus*, se detectaron 2.054 péptidos exclusivos en *M. musculus* y 3040 en *M. spretus*. La mayoría de los péptidos exclusivos mostraron un número reducido de espectros, generalmente menor a 10, sugiriendo que la baja abundancia de los péptidos puede ser la causa principal por la que no fueron detectados en algunas especies. De hecho, cuando la distribución del número de espectros de los péptidos con detección exclusiva fue comparada con la de los péptidos detectados en ambas especies, los primeros mostraron una media significativamente menor (Figura suplementaria 9A). Se diseñó entonces una tabla de contingencia con el fin de comparar la frecuencia de péptidos exclusivos y compartidos entre las proteínas con diferencias de abundancia significativas y no significativas. Un test chi-cuadrado no encontró diferencias significativas entre las categorías comparadas (*M. musculus* $p = 1$, *M. spretus* $p = 1$) (Figura suplementaria 9A). Estos resultados por tanto indican que los péptidos con detección exclusiva no contribuyen a la tasa de falsos positivos, apoyando la robustez de los análisis comparativos. Los mismos resultados fueron obtenidos para las comparaciones *M. musculus*-*M. spicilegus* (Figura suplementaria 9B) y *M. spretus*-*M. spicilegus* (Figura suplementaria 9C).

4.28. *Cambios en abundancia de proteínas dirigidos por la competición espermática (análisis APEX)*

Con el fin de identificar proteínas que pueden experimentar cambios en abundancia dirigidos por selección sexual, se llevaron a cabo comparaciones simultáneas de los valores de abundancia estimados para *M. musculus*, que es la especie con niveles bajos de competición espermática, con los valores estimados para *M. spretus* y *M. spicilegus*, que son especies mostrando niveles altos de competición espermática. Se identificaron 9 proteínas mostrando mayor abundancia en *M. musculus* respecto a las otras dos especies (Tabla XXI). Seis de estas proteínas tienen funciones biológicas anotadas. Además, las proteínas Akap4, Gpd2 y Gpx4 se han asociado con fenotipos reproductivos. Ratones mutantes para los genes *Akap4* y *Gpx4* se relacionan con fenotipos de morfología anormal del flagelo y con motilidad reducida, mientras que el fenotipo mutante de *Gpd2* presenta fertilidad reducida (Tabla suplementaria 10). La proteína Spry presenta un papel en el transporte de iones. La peroxidasa Hnrpl1 está implicada en la respuesta a especies reactivas de oxígeno. La fosfofructokinasa Pfkfb3 participa en el metabolismo de la glucosa, y un miembro de la familia clusterina se ha asociado con funciones múltiples tales como morfogénesis celular, apoptosis o respuesta al estrés oxidativo. Por otro lado, ocho proteínas mostraron

una mayor abundancia de forma significativa en las especies *M. spretus* y *M. spicilegus* (Tabla XXI). Entre estas, siete proteínas se han asociado con procesos biológicos. La zonadhesina (*Zan*) y la proteína de unión a zona pelúcida 2 (*Zpbp2*) tienen fenotipos relacionados con el reconocimiento espermatozoide-óvulo. El mutante de *Zan* también presenta espermatozoides con una fisiología anormal, mientras que en el mutante para *Zpbp2* se han identificado diversos fenotipos que incluyen desde fertilidad reducida hasta morfología o fisiología anormal del espermatozoide (Tabla suplementaria 10). *Dnahc3* es una proteína asociada con los microtúbulos del flagelo. *Man2b2* está implicada en el metabolismo de la manosa, y *Mecr* es una proteína mitocondrial que participa en el metabolismo de ácidos grasos. La proteína ADAM4 presumiblemente tiene un papel proteolítico.

Tabla XXI. Lista de proteínas con diferencias de abundancia entre las especies de baja competición espermática (*M. musculus*) y alta competición espermática (*Mus spretus* y *Mus spicilegus*).

Proteína	Nombre de la proteína	Cociente APEX		Z-score		<i>p</i>	
		<i>M. musculus</i>	<i>M. musculus</i>	<i>M. musculus</i>	<i>M. musculus</i>	<i>M. musculus</i>	<i>M. musculus</i>
		vs <i>M. spretus</i>	vs <i>M. spicilegus</i>	vs <i>M. spretus</i>	vs <i>M. spicilegus</i>	vs <i>M. spretus</i>	vs <i>M. spicilegus</i>
BSPRY	B box and SPRY domain-containing protein	2,396	2,322	4,429	4,226	9,49E-06	2,38E-05
AKAP4	A-kinase anchor protein 4 isoform b	1,126	2,410	4,831	14,813	1,36E-06	0,00E+00
GPD2	glycerol-3-phosphate dehydrogenase, mitochondrial precursor	1,509	2,257	4,981	8,698	6,33E-07	0,00E+00
PFKM	6-phosphofructokinase, muscle type	2,573	3,003	5,258	5,623	1,45E-07	1,88E-08
HNRPLL	phospholipid hydroperoxide glutathione peroxidase, nuclear isoform 1	2,014	2,307	5,502	5,807	3,76E-08	6,38E-09
GSTM5	glutathione S-transferase Mu 5	1,642	1,706	5,875	4,876	4,22E-09	1,08E-06
PACRG	parkin coregulated gene protein homolog	4,191	5,004	6,365	6,488	1,96E-10	8,71E-11
IRGC1	interferon-inducible GTPase 5	2,920	2,214	8,478	6,581	0	4,68545E-11
LOC100046120	clusterin precursor	5,363	4,530	10,018	8,957	0	0
ZAN	zonadhesin precursor	2,829	2,522	-13,052	-11,916	6,23E-39	9,75E-33
ZBPB2	zona pellucida-binding protein 2 isoform 1 precursor	3,226	2,483	-6,750	-4,743	1,48E-11	2,10E-06
DNAHC3	dynein heavy chain 3, axonemal	4,068	3,346	-5,940	-4,965	2,85E-09	6,87E-07
MECR	trans-2-enoyl-CoA reductase, mitochondrial precursor	6,759	7,827	-5,396	-6,021	6,82E-08	1,73E-09
CD109	CD109 antigen precursor	2,753	2,939	-4,864	-5,111	1,15E-06	3,20E-07
MAN2B2	epididymis-specific alpha-mannosidase precursor	2,408	2,759	-4,861	-5,545	1,17E-06	2,93E-08
ADAM4b	a disintegrin and metalloprotease domain 4b precursor	10,551	10,752	-4,815	-4,866	1,47E-06	1,14E-06
Gm884	uncharacterized protein LOC380730	1,762	1,769	-4,598	-4,510	4,27E-06	6,49E-06

Las proteínas con mayor abundancia en el proteoma de *M. musculus* se representan por encima de la línea discontinua, mientras que la proteínas con mayor abundancia en *Mus spretus* y *Mus spicilegus* se representan debajo de la línea.

4.29. Análisis comparativos en geles 2D (2D-DIGE)

Se prepararon geles en 2 dimensiones (2D) en los que se separaron las proteínas del espermatozoide de cada especie de ratón. Entre las 2 réplicas de geles que se prepararon para cada especie, se identificaron un total de 1.198 *spots* para *M. musculus*, 1.186 *spots* para *M. spretus* y 1.175 *spots* para *M. spicilegus*. Entre estos, 739 *spots* coincidieron entre *M. musculus* y *M. spretus*, 838 entre *M. musculus* y *M. spicilegus*, y 911 entre *M. spretus* y *M. spicilegus*. Se calculó un cociente para comparar la señal de fluorescencia entre los *spots* marcados con diferentes sondas. Se detectaron un total 250 *spots* mostrando diferencias significativas en la señal de fluorescencia (lo que equivale al volumen del *spot*) entre *M. musculus* y *M. spretus*, 253 entre *M. musculus* y *M. spicilegus*, y 268 entre *M. spretus* y *M. spicilegus*. Tras la edición manual para eliminar cualquier artefacto que genere una señal y separar las cadenas formadas por varios *spots*, así como una validación visual de la coincidencia de los *spots* entre los geles comparados, se seleccionaron 29, 34 y 33 *spots* significativos de cada comparación para su identificación por espectrometría de masas (Fig. 36). El bajo poder de resolución de los geles 2D en los rangos de pH y masa conllevó a la identificación de múltiples proteínas dentro de cada *spot*. Por otro lado, numerosas proteínas fueron identificadas en más de un *spot*, debido principalmente a modificaciones post-traduccionales que experimentan muchas proteínas y a la generación de diversas isoformas por *splicing*. Los datos de MS por tanto se procesaron manualmente siguiendo los siguientes pasos: 1) descarte de las proteínas identificadas por un único péptido; 2) selección solo de aquellas proteínas cuyos péptidos identificados cubren más del 25 % de la secuencia; 3) ordenamiento de las proteínas en función del número de péptidos identificados. Tras aplicar estos criterios, se obtuvieron un total de 511 identificaciones para 177 proteínas (Tabla suplementaria 15). Las proteínas más predominantes fueron, de acuerdo con los resultados de los análisis de LC-MS/MS, enzimas metabólicas implicadas en la glicólisis y proteínas catalíticas localizadas en la mitocondria.

Con el fin de encontrar qué proteínas generan las diferencias en abundancia detectadas entre los *spots*, se trató de identificar la proteína más candidata dentro de cada *spot* en base a los siguientes parámetros: 1) un elevado número de péptidos identificados; 2) un elevado porcentaje de cobertura; y 3) una concordancia con la posición esperada en el gel 2D de acuerdo con su masa molecular, su punto isoelectrico, y la información recogida de la base de datos de geles 2D de reproducción (<http://reprod.njmu.edu.cn/cgi-bin/2d>). De este modo, se mapearon un total de 34 *spots* asociados con una sola proteína a lo largo de las tres comparaciones interespecíficas (Fig. 36, Tabla suplementaria 15). Algunas de las proteínas seleccionadas coincidieron con más de un *spot* (ej: Sord, Aldoa y Aco2).

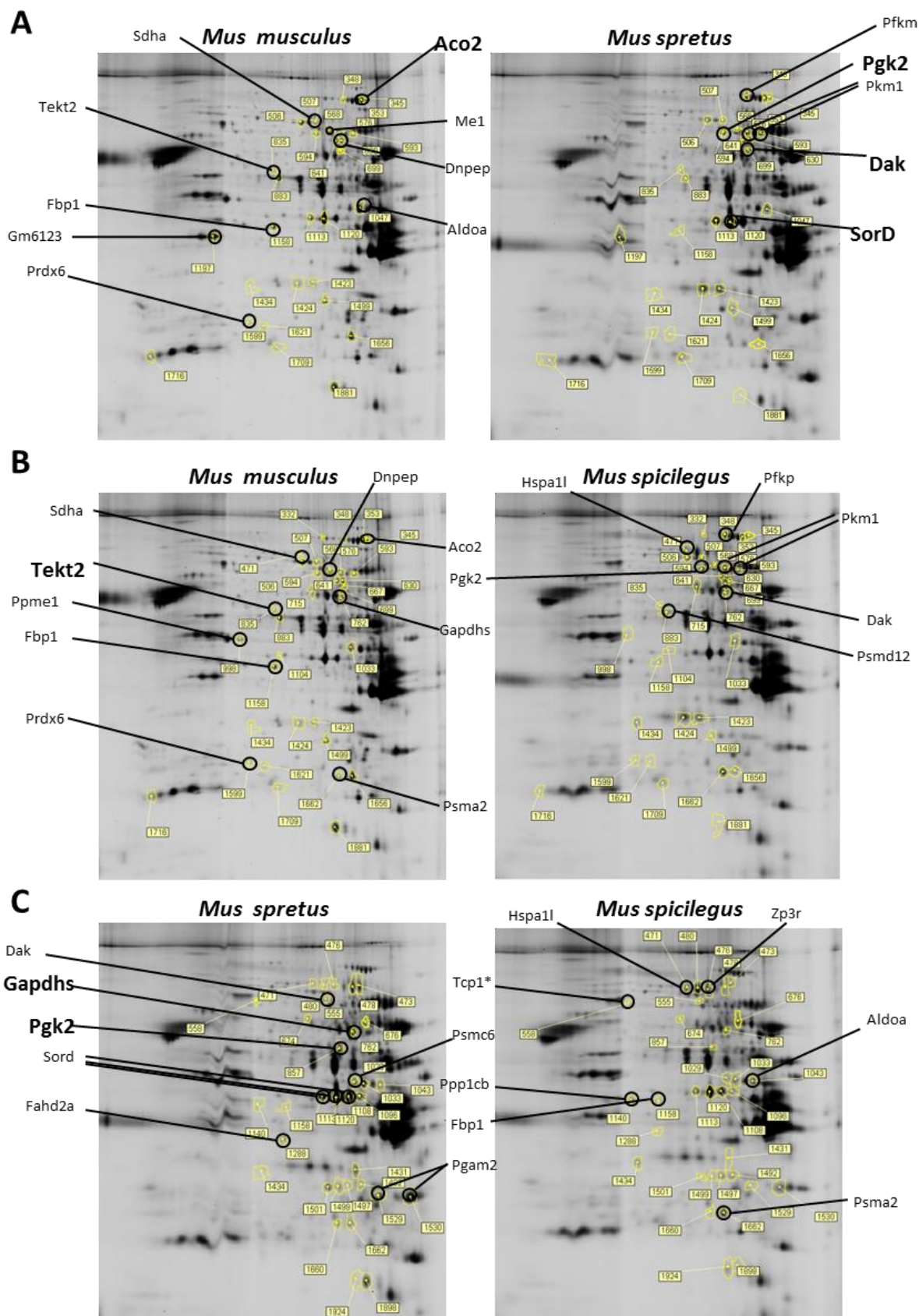


Figura 36. Comparación de geles 2D entre especies de roedores. Los spots mostrando diferencias en abundancia están etiquetados con su número de referencia en el gel. Las proteínas asociadas con un único spot se señalan dentro del gel. Las proteínas que muestran diferencias cuantitativas entre los análisis en geles 2D (DIGE) y los análisis de LC-MS/MS (APEX) se representan en negrita. (*) La proteína Tcp1 mostró un patrón cuantitativo opuesto entre las dos metodologías.

4.30. Cambios en abundancia de proteínas dirigidos por la competición espermática (Análisis 2D-DIGE)

Como se describió para los análisis de APEX, se identificaron proteínas con diferencias significativas en abundancia entre las especies de baja y alta competición espermática. Entre las proteínas que muestran mayor abundancia en *M. musculus*, solo Tekt2, un componente estructural del flagelo del espermatozoide, presenta un fenotipo reproductivo asociado con una morfología anormal del flagelo e infertilidad (Tabla suplementaria 10). Dentro de este grupo se encontraron las enzimas mitocondriales Aco2, implicada en la isomerización de citrato, Sdha, que participa en la cadena transportadora de electrones, y la enzima málica Me1. Las enzimas Fbp1, implicada en gluconeogénesis, y la aminopeptidasa Dnpep también mostraron mayor abundancia en *M. musculus* (Tabla XXII). Para las especies de alta competición espermática (*Mus spretus* y *Mus spicilegus*), se detectó una mayor abundancia para las enzimas glicolíticas Pfkp, Pkg2 y dos isoformas de la piruvato kinasa. La proteína Pkg2 se ha asociado con un fenotipo de infertilidad masculina (Tabla suplementaria 10). La enzima DAK, implicada en metabolismo de glicerol, también mostró este patrón (Tabla XXII).

Tabla XXII. Proteínas mostrando diferencias en abundancia en *M. musculus* respecto a las especies *Mus spretus* y *Mus spicilegus* en los análisis de geles 2D.

Proteína	Nombre de la proteína	Número del spot en gel 2D	Cociente promedio		Test-t	
			<i>M. musculus</i> vs <i>M. spretus</i>	<i>M. musculus</i> vs <i>M. spicilegus</i>	<i>M. musculus</i> vs <i>M. spretus</i>	<i>M. musculus</i> vs <i>M. spicilegus</i>
ACO2	aconitate hydratase mitochondrial precursor	345	-4,47	-3,27	0,00011	0,00022
ACO2	aconitate hydratase mitochondrial precursor	353	-3,57	-3,1	0,0095	0,011
SDHA	succinate dehydrogenase [ubiquinone] flavoprotein subunit mitochondrial precursor	507	-1,37	-1,61	0,054	0,0032
ME1	NADP-dependent malic enzyme isoform 1	568	-3,47	-3,15	0,0034	0,0094
DNPEP	aspartyl aminopeptidase isoform b	641	-1,72	-2,19	0,042	0,0013
TEKT2	tektin-2	835	-4,35	-5,2	0,007	0,0051
FBP1	fructose-1, 6-bisphosphatase 1	1158	-18,3	-4,76	0,0067	0,00059
PFKP	6-phosphofructokinase type C	348	2,29	2,33	0,011	0,011
PKM	pyruvate kinase isoform 2	576	1,77	1,65	0,005	0,000026
Gm6560	pyruvate kinase isoform 1	593	2,02	1,66	0,001	0,011
PGK2	phosphoglycerate kinase 2	594	1,89	1,87	0,018	0,018
DK	bifunctional ATP-dependent dihydroxyacetone kinase	699	1,51	2,34	0,0025	0,0016

Las proteínas con mayor abundancia en *M. musculus* se representan por encima de la línea discontinua, mientras que aquellas con mayor abundancia en *Mus spretus* y *Mus spicilegus* se muestran por debajo de la línea.

4.31. Comparación de los datos obtenidos entre los análisis de APEX y 2D-DIGE

De las 50 proteínas que mostraron diferencias cuantitativas entre las especies de *M. musculus* y *Mus spretus* en los análisis de APEX, cuatro se identificaron también entre los *spots* seleccionados dentro de los geles 2D (Tabla XXIII, Fig. 36A). En dos casos (Aco2 y Sord), las proteínas significativas de APEX fueron identificadas en más de un *spot* de los geles. Las diferencias cuantitativas detectadas por los análisis de APEX fueron concordantes con aquellas detectadas en los geles 2D en todos los casos (Tabla XXIII). El *spot* asignado a la proteína Aco2 mostró mayor intensidad en *M. musculus*, mientras que los *spots* donde se encuentran las enzimas Dak, Sord y Pgk2 presentaron mayor abundancia en *M. spretus*.

Solamente una de las proteínas mostrando diferencias significativas entre *M. musculus* y *M. spicilegus* también fue identificada en los análisis 2D-DIGE: la tektina Tekt2 (Tabla XXIII, Fig. 36B). Esta proteína mostró mayor abundancia en los geles de *M. musculus*, apoyando los datos comparativos de APEX (Tabla XXIII). Tres de las 33 proteínas con abundancia diferencial entre *M. spretus* y *M. spicilegus* se identificaron dentro de los *spots* de geles 2D (Tabla XXIII, Fig. 36C). Las enzimas metabólicas Pgk2 y Gapdhs se asociaron robustamente, con *spots* mostrando mayor abundancia en el gel de *M. spretus*, coincidiendo así con los resultados de APEX. No obstante, la proteína Tcp1 mostró un patrón de abundancia opuesto entre los análisis de APEX y 2D-DIGE.

Tabla XXIII. Datos cuantitativos de proteínas identificadas en los análisis de LC-MS/MS-APEX y 2D-DIGE.

Proteína	Nombre de la proteína	Número de <i>spot</i> en el gel 2D (cociente promedio) ^a	Patrón 2D-DIGE	Z-score de APEX ^b	Patrón de APEX
<i>M. musculus</i> vs <i>Mus spretus</i>					
ACO2	aconitate hydratase mitochondrial precursor	345 (-4,47), 353 (-3,1)	<i>musculus</i> > <i>spretus</i>	4,78	<i>musculus</i> > <i>spretus</i>
PGK2	phosphoglycerate kinase 2	594 (1,87)	<i>spretus</i> > <i>musculus</i>	-6,17	<i>spretus</i> > <i>musculus</i>
DAK	bifunctional ATP-dependent dihydroxyacetone kinase/FAD-AMP lyase	699 (2,34)	<i>spretus</i> > <i>musculus</i>	-6,40	<i>spretus</i> > <i>musculus</i>
SORD	sorbitol dehydrogenase	1113 (-2,48), 1120 (2,16)	<i>spretus</i> > <i>musculus</i>	-8,91	<i>spretus</i> > <i>musculus</i>
<i>M. musculus</i> vs <i>Mus spicilegus</i>					
TEKT2	tektin-2	835 (-4,0)	<i>musculus</i> > <i>spicilegus</i>	4,3	<i>musculus</i> > <i>spicilegus</i>
<i>Mus spretus</i> vs <i>Mus spicilegus</i>					
TCP1	T-complex protein 1 subunit alpha	558 (2,0)	<i>spicilegus</i> > <i>spretus</i>	7,08	<i>spretus</i> > <i>spicilegus</i>
GAPDHS	glyceraldehyde-3-phosphate dehydrogenase testis-specific	762 (-2,0)	<i>spretus</i> > <i>spicilegus</i>	6,18	<i>spretus</i> > <i>spicilegus</i>
PGK2	phosphoglycerate kinase 2	857 (-2,65)	<i>spretus</i> > <i>spicilegus</i>	6,1	<i>spretus</i> > <i>spicilegus</i>

^aLos cocientes promedio fueron calculados usando los valores de *Mus spretus* y *Mus spicilegus* como geles de referencia en las comparaciones respectivas con los geles de *M. musculus*. La especie *Mus spicilegus* fue usada como referencia en la comparación *Mus spretus* vs *Mus spicilegus*.

^bLos Z-scores de APEX se calcularon usando *M. musculus* como especie de referencia en las comparaciones respectivas con *Mus spretus* y *Mus spicilegus*. La especie *Mus spicilegus* fue usada como referencia en la comparación *Mus spretus* vs *Mus spicilegus*.

5. DISCUSIÓN

El objetivo de esta Tesis Doctoral fue analizar la evolución de las proteínas reproductivas que se expresan en espermatozoides de roedores e identificar proteínas que experimentan una evolución adaptativa dirigida por fuerzas de selección sexual, concretamente la competición espermática. Para ello, se desarrollaron estudios con diferentes enfoques: a) estudios de genes candidatos en los que se aplicaron análisis comparativos entre especies para evaluar la influencia de la competición espermática sobre la evolución de una proteína en cuestión, y b) estudios globales en los que se analizó la evolución del proteoma del espermatozoide con el fin de detectar grupos de proteínas más influenciados por las fuerzas selectivas. Por otra parte, con estos estudios se trató de analizar diferentes formas de evolución, tratando de demostrar que los genes reproductivos no sólo evolucionan adaptativamente a nivel de secuencia codificante, sino que las fuerzas selectivas también pueden promover cambios en la estructura y la expresión de las proteínas.

En los diferentes estudios que se presentan se han utilizados grupos de roedores pertenecientes a las familias Muridae y Cricetidae. El uso de estas especies de roedores provee un modelo adecuado para estudios de selección sexual porque, a pesar de su proximidad filogenética, muestran una amplia variedad en sus sistemas de apareamiento y, por lo tanto, en sus niveles de competición espermática (Dewsbury 1984; Breed y Taylor 2000; Gómez Montoto *et al.* 2011). Esto permite minimizar los efectos filogenéticos y atribuir las diferencias inter-específicas en los rasgos reproductivos a diferencias en la intensidad de selección sexual. Además, el uso de especies cercanas provee una ventaja para estudios de evolución molecular, ya que permite detectar los primeros cambios que se producen en las proteínas tras la especiación e incluso se pueden identificar mutaciones que han podido contribuir al aislamiento reproductivo y a la formación de especies (Ramm *et al.* 2008, 2009; Martín-Coello *et al.* 2009; Lüke *et al.* 2011).

En un primer estudio se analizó la evolución molecular de proteínas implicadas en los diferentes procesos que ocurren durante la vida del espermatozoide. El objetivo de este estudio fue comprobar si las fuerzas evolutivas actúan con diferente intensidad entre los diferentes procesos que tienen lugar en el espermatozoide desde su formación hasta el momento de fecundación, e identificar los procesos más impactados por la selección positiva. Los resultados de estos análisis revelaron que las proteínas implicadas en la interacción de gametos evolucionan más rápidamente que el resto de las proteínas del espermatozoide. Además, una alta proporción de estas proteínas evolucionan bajo selección positiva. Nuestros resultados por tanto sugieren que los mecanismos de interacción entre gametos pueden ser una diana importante para las fuerzas de selección sexual, lo que coincide con estudios previos que han encontrado evidencia de selección positiva en numerosas proteínas implicadas en fecundación (Torgerson *et al.* 2002; Civetta 2003a; Swanson *et al.* 2003; Dorus *et al.* 2010).

La rápida divergencia de las proteínas implicadas en la interacción espermatozoide-óvulo de roedores puede estar asociada a procesos de coevolución por los que las, proteínas en la superficie del espermatozoide y del óvulo que se reconocen y unen durante la fecundación, experimentan cambios constantes para mantener la interacción entre los gametos, como ha sido descrito en invertebrados marinos (Metz y Palumbi 1996; Swanson y Vacquier 2002; Galindo *et al.* 2003; Vacquier y Swanson 2011). Tal proceso de coevolución va a ser el responsable de mantener interacciones específicas entre los gametos de cada especie, por lo que puede ser importante para el aislamiento reproductivo y la especiación. Sobre la fuerza selectiva que promueve la divergencia de proteínas de fecundación, se ha sugerido en invertebrados que, en especies con intensa competición espermática, el conflicto sexual que se genera va a seleccionar genotipos raros en las proteínas del óvulo que reduzcan la afinidad de unión de los espermatozoides para así evitar la polispermia (Gavrilets 2000; Levitan y Ferrell 2006; Levitan *et al.* 2007). En esta situación, sólo los espermatozoides con genotipos raros tendrían mayor oportunidad de unirse al óvulo y la competición espermática incrementará estos genotipos en la población, repitiéndose nuevamente el ciclo (Swanson y Vacquier 2002). En este contexto, nuestro trabajo plantea la posibilidad de que, en especies con fecundación interna, donde los gametos están expuestos a un mayor número de factores regulando la fecundación, las interacciones espermatozoide-óvulo pueden tener también un papel clave en la selección sexual y el aislamiento reproductivo.

Por otro lado, nuestros análisis detectaron una alta proporción de proteínas con función en la motilidad del espermatozoide mostrando evidencia de selección positiva. Se ha demostrado en diversos estudios que la competición espermática favorece un incremento en la velocidad de natación de los espermatozoides mediante adaptaciones morfológicas y fisiológicas (Gomendio y Roldan 2008; Fitzpatrick *et al.* 2009; Lüpold *et al.* 2009; Gómez Montoto *et al.* 2011b; Tourmente *et al.* 2011). En este contexto, nuestros resultados sugieren que la competición espermática también podría dirigir la evolución adaptativa de proteínas implicadas en la motilidad espermática para lograr un incremento en la velocidad de natación de los espermatozoides.

Se debe tener en cuenta que, a pesar de que se asume la importancia de la competición espermática en la evolución de las proteínas del espermatozoide, las cepas analizadas en este estudio no han evolucionado bajo la influencia de la selección sexual. Para evaluar el impacto de la competición espermática sobre la evolución de los componentes del espermatozoide serían necesarios estudios evolutivos comparando especies con diferentes niveles de promiscuidad.

Las proteínas implicadas en otros procesos diferentes a la interacción espermatozoide-óvulo y a la motilidad espermática exhibieron evolución adaptativa en una menor proporción. No obstante, este hecho no implica que procesos como la espermatogénesis, la capacitación o la reacción acrosómica estén menos expuestos a la acción de las fuerzas selectivas, ya que este y otros estudios

(Wyckoff *et al.* 2000; Torgerson *et al.* 2002; Dorus *et al.* 2010) han mostrado evidencia de evolución adaptativa en diversas proteínas asociadas con estos procesos.

Estudios previos en *Drosophila* y ratón han propuesto la hipótesis de que el espermatozoide es una célula sujeta a selección diferencial, resultando en una compartimentación de la adaptación entre los diferentes componentes subcelulares (Dorus *et al.* 2006, 2010). Los resultados del presente estudio indican que también puede existir una compartimentación de la adaptación en términos del papel que desempeñan las proteínas en el espermatozoide. Por lo tanto, los resultados de este trabajo y de estudios previos sugieren que en el espermatozoide existe una compartimentación tanto a nivel subcelular como funcional, probablemente como una respuesta a la acción diferencial de las múltiples fuerzas selectivas que dirigen la evolución del espermatozoide.

Otra aportación importante de este trabajo es que se han identificado un gran número de proteínas divergiendo rápidamente que podrían ser el centro atención para futuros estudios evolutivos y funcionales. Además, la identificación de aminoácidos sujetos a selección positiva en un amplio grupo de proteínas reproductivas puede dar indicaciones sobre regiones proteicas importantes para la función del espermatozoide.

En cualquier caso, el uso de cepas y especies estrechamente relacionadas de ratón puede proveer un modelo adecuado para el desarrollo de análisis evolutivos (Anisimova *et al.* 2001), así como una herramienta útil para el diseño de estudios que evalúen los efectos de la variación genómica (Keane *et al.* 2011).

En el segundo estudio, se analizó la evolución de la proteína Catsper1, una subunidad que forma parte del canal de calcio Catsper, que está implicado en la motilidad espermática y es esencial para la hiperactivación del espermatozoide (Ren *et al.* 2001; Carlson *et al.* 2005; Qi *et al.* 2006; Liu *et al.* 2007). Se analizó la región amino-terminal de Catsper1, que es un dominio intracelular sobre el que se ha observado una alta divergencia tanto a nivel de secuencia como en longitud (Podlaha y Zhang 2003; Podlaha *et al.* 2005). Dado que Catsper1 regula la motilidad del espermatozoide y este rasgo es importante para el éxito reproductivo, se evaluó si la divergencia de esta proteína está relacionada con los niveles de competición espermática. Mientras que la tasa de evolución molecular de la secuencia de Catsper1 no mostró una relación con los niveles de competición espermática, sí se detectó una correlación negativa entre el tamaño de la secuencia y la intensidad de competición espermática. Este hallazgo da evidencia de que la competición espermática es capaz de influir sobre la dinámica estructural de las proteínas reproductivas. Previamente solo un estudio había demostrado una asociación entre la masa molecular de la proteína SvsII, un componente del fluido seminal, con la intensidad de competición espermática en roedores (Ramm *et al.* 2009).

En adición, se observó que el acortamiento de la región amino terminal de *Catsper1* se relaciona con una mayor velocidad de natación del espermatozoide. Estos resultados por tanto apoyan la hipótesis formulada en el primer estudio de que la competición espermática es capaz de favorecer un incremento en la velocidad de natación del espermatozoide a través de cambios adaptativos en las proteínas implicadas en la regulación de la motilidad. En cualquier caso, este sería el primer estudio en el que se establece una asociación entre la evolución estructural de una proteína reproductiva y la adaptación de un rasgo fenotípico.

En conjunto, basados en la correlación observada entre la masa relativa de testículo, la secuencia de *Catsper1* y los parámetros de velocidad espermática, nuestros resultados apoyan la hipótesis de que la competición espermática favorece un acortamiento del dominio intracelular de *Catsper1*, que implica un incremento en la velocidad de natación del espermatozoide durante su transporte en el tracto reproductivo de la hembra, teniendo así mayor probabilidad de éxito en la fecundación.

Aunque los mecanismos que generan inserciones y deleciones no se comprenden en profundidad, se ha observado que las pequeñas deleciones se generan con mayor frecuencia que las inserciones de tamaño similar en las secuencias codificantes de proteína (de Jong y Rydén 1981). Nuestros datos para *Catsper1* concuerdan con esta tendencia. No obstante, en las ramas terminales de la filogenia se observó que algunos linajes muestran una relación inserción: deleción que se desvía de la esperada. Los casos más extremos fueron las especies promiscuas *Apodemus sylvaticus* y *Lemniscomys barbarus*, que mostraron casi una presencia exclusiva de deleciones. Por tanto, parece que la mayor fijación de deleciones en el primer exón del gen de *Catsper1* no sería solo resultado del sesgo impuesto por el mecanismo de los indels, sino que otras fuerzas selectivas como la competición espermática se encontrarían favoreciendo el acortamiento de esta región.

A pesar de que se encontró un elevado número de mutaciones adaptativas en la región N-terminal de *Catsper1*, no se encontró evidencia de una asociación de los valores de ω con los niveles de competición espermática. Una posible interpretación es que, mientras los indels parecen ser la diana primaria de la selección sexual post-cópula, la divergencia de aminoácidos puede estar influida por múltiples fuerzas selectivas que promueven cambios no relacionados con el éxito reproductivo. Por otra parte, es posible que las diferentes especies adquieran diferentes rutas de mutaciones para incrementar el éxito reproductivo, conllevando a una evolución adaptativa heterogénea entre linajes. Por lo tanto, el efecto de la selección sexual sobre la evolución molecular de *Catsper1* no se puede descartar a pesar de que los resultados no revelaron tal asociación.

Actualmente la función del dominio intracelular de *Catsper1* no ha sido caracterizada, aunque se ha sugerido que actúa como un sensor de pH debido a su riqueza en histidinas (Navarro *et al.* 2008). No obstante, nuestros análisis no revelaron una asociación entre la abundancia de este

aminoácido y los niveles de competición espermática, así como con los parámetros de velocidad de los espermatozoides. En su lugar, se observó que diferentes clados de especies de múridos muestran una cantidad constante de histidinas, lo que sugiere que las ganancias (o pérdidas) de este residuo han ocurrido de acuerdo con el tiempo evolutivo. No obstante, los cambios de histidinas en Catsper1 aún podrían ser adaptativos, ya que casi la mitad de aminoácidos bajo selección positiva implican a estos residuos. Por otra parte, trabajos previos han revelado que la sensibilidad de los canales de potasio al pH intracelular es regulada por histidinas localizadas en posiciones particulares dentro de un dominio intracelular (Coulter *et al.* 1995; Chanchevalap 2000). Por tanto, es posible que la sensibilidad de Catsper1 a pH dependa más de la posición de las histidinas que de su abundancia.

Otro modelo funcional propone que la región N-terminal de Catsper1 actuaría como una compuerta física que bloquea el canal Catsper, inactivándolo (Podlaha y Zhang 2003). Bajo este modelo, la variación en longitud del dominio intracelular tendría implicaciones importantes en la tasa de activación/inactivación del canal debido a restricciones de espacio (Hoshi *et al.* 1990), y esto podría tener efectos sobre la motilidad espermática.

Independientemente de la función molecular que desempeña el dominio amino-terminal de Catsper1, la actividad del canal parece verse afectada significativamente por los cambios en longitud de esta región, y probablemente esta sea la razón por la que las inserciones y deleciones en esta región se encuentran favorecidas por selección positiva.

Este estudio constituye la primera evidencia de cómo la competición espermática es capaz de influir en rasgos importantes para el éxito reproductivo a través de cambios estructurales en una proteína del espermatozoide. Las proteínas reproductivas experimentan cambios de longitud mediante indels con mayor frecuencia que las proteínas expresadas en otros tejidos (Torgerson *et al.* 2002). Siempre y cuando los indels no impliquen cambios drásticos para la función de la proteína, la divergencia estructural puede ser una fuente de variación capaz de promover cambios ventajosos de manera más eficiente que las mutaciones de nucleótidos. Por lo tanto, es posible que los indels constituyan una diana primaria de selección positiva en muchas proteínas reproductivas.

Además de promover incrementos en el número y la velocidad de natación de los espermatozoides, se ha comprobado que la competición espermática también es capaz de influir sobre procesos que ocurren cercanos al momento de la fecundación. Por tanto, el objetivo de otro estudio fue el de examinar si la competición espermática influye sobre la evolución de PKDREJ, una proteína de la superficie del espermatozoide supuestamente implicada en la regulación de la reacción acrosómica (Sutton *et al.* 2006; Hamm *et al.* 2007). Estudios con ratones transgénicos han encontrado que la pérdida del gen *Pkdrej* genera espermatozoides fértiles pero con menor éxito reproductivo debido a deficiencias en el transporte y un retraso en el inicio de la reacción acrosómica (Sutton *et al.* 2008). PKDREJ por tanto se propuso como un factor importante para la selección sexual post-

cópula. Los análisis evolutivos de PKDREJ, particularmente sobre el dominio extracelular REJ, fueron llevados a cabo en dos familias de roedores: Muridae y Cricetidae. El dominio REJ es una región que supuestamente actúa como receptor de otras células y señales extracelulares (Moy *et al.* 1996; Delmas 2005). Entre las especies de la familia Muridae no se encontró evidencia de evolución adaptativa. La ausencia de una señal de selección positiva en la familia Muridae probablemente se debe a que estas especies son muy próximas filogenéticamente y por tanto presentan secuencias del gen *Pkdrej* muy conservadas. Este fenómeno ya se ha observado previamente cuando se han analizado otras proteínas conservadas en roedores como la protamina 1 (Prm1) (Martín-Coello *et al.* 2009). No obstante, no se puede descartar la posibilidad de que PKDREJ evoluciona libre de la influencia de fuerzas selectivas en este grupo de roedores. Por otra parte, se detectó un par de aminoácidos bajo selección positiva en las especies de la familia Cricetidae. El elevado nivel de conservación del dominio REJ en roedores contrasta con la intensa selección positiva encontrada para esta misma región en primates y erizos de mar (Mah *et al.* 2005; Hamm *et al.* 2007), sugiriendo que la proteína PKDREJ se encuentre sujeta a una mayor selección purificante en roedores.

Curiosamente, la débil señal de selección positiva detectada en PKDREJ de cricétidos se concentra en las especies con los niveles más elevados de competición espermática, como son *Cricetulus griseus* y *Mesocricetus auratus*. Además, se observó una correlación positiva entre los niveles de competición espermática y la tasa evolutiva de PKDREJ. Estos resultados indican que, a pesar del alto grado de conservación de PKDREJ en roedores, los pocos cambios adaptativos que ocurren en esta proteína se fijan en las especies más promiscuas, por lo que estos cambios pueden contribuir a incrementar el éxito reproductivo de algún modo. La función molecular de PKDREJ en el espermatozoide no se ha caracterizado enteramente y por tanto no se puede inferir el significado funcional de los cambios adaptativos. Se ha sugerido que PKDREJ participa en el desencadenamiento de la reacción acrosómica en respuesta a la zona pelúcida (Sutton *et al.* 2006; Hamm *et al.* 2007). En este sentido, las mutaciones adaptativas que ocurran en el dominio REJ podrían afectar a la eficiencia con la que PKDREJ responde a las señales de la zona pelúcida, influyendo en el éxito de la fecundación. Por otro lado, el fenotipo observado en ratones transgénicos hace pensar que PKDREJ pueda estar implicado en alguna clase de interacción espermatozoide-oviducto, donde entran en juego fuerzas de selección post-cópula como la selección críptica de la hembra o la propia competición espermática. En este escenario, la producción de cambios adaptativos en PKDREJ puede afectar al comportamiento del espermatozoide en el oviducto y por tanto influir sobre la eficiencia de fecundación.

Dado que otros dominios de PKDREJ diferentes del REJ probablemente son importantes para la función del espermatozoide, sería interesante analizar la evolución de la secuencia completa de PKDREJ. Esto podría revelar, por ejemplo, si los segmentos transmembrana de PKDEJ, que podrían

tener una función análoga a la proteína PKD-2 como un canal permeable de calcio (Delmas 2005), muestran señales de evolución adaptativa.

En el caso de primates, un estudio previo reveló que la proteína PKDREJ muestra una intensa señal de selección positiva, pero no se detectó una asociación entre los episodios de evolución adaptativa de PKDREJ y los niveles de competición espermática (Hamm *et al.* 2007). La ausencia de esta correlación se atribuyó en un primer momento a que el índice de competición espermática utilizado en este trabajo se basó en los diferentes sistemas de apareamiento descritos en primates, que son medidas normalmente imprecisas o que presentan limitaciones (Hamm *et al.* 2007).

La divergencia de PKDREJ tampoco se asoció con la masa relativa de testículo, dando mayor evidencia de que la competición espermática no es una fuerza predominante dirigiendo la evolución de PKDREJ en primates. No obstante, esto no implica necesariamente que PKDREJ evolucione al margen de la influencia de las fuerzas de selección sexual. El hecho de detectar selección positiva en esta proteína con expresión específica en testículo y con un papel importante en el espermatozoide indica que PKDREJ probablemente experimente cambios dirigidos por fuerzas selectivas relacionadas con el sexo. Una posible explicación para la ausencia de correlación es que diversas fuerzas selectivas pueden participar conjuntamente en la evolución de PKDREJ. En este escenario, aquellos cambios adaptativos fijados específicamente por selección sexual se enmascararían por los cambios fijados por otras fuerzas de cara a los análisis evolutivos. Esta teoría sería apoyada por el hecho de que el dominio REJ de primates muestra una fuerte señal de selección positiva (Hamm *et al.* 2007).

Como se mencionó anteriormente, una característica esencial de la reproducción es que las proteínas implicadas en el reconocimiento de los gametos mantengan su interacción en las especies. Para mantener esta interacción, debe darse un proceso de coevolución por el que los cambios que se producen en la proteína de uno de los gametos se compensan por cambios en la proteína del otro gameto (Swanson y Vacquier 2002). Dado que la coevolución va a ser un proceso característico de las proteínas que interaccionan directamente durante la fecundación, se pueden utilizar las señales de coevolución como herramienta para predecir interacciones entre pares de proteínas del espermatozoide y el óvulo (Clark *et al.* 2009). Por lo tanto, otro estudio consistió en analizar la evolución de dos proteínas esenciales en la fusión de gametos en ratón (Izumo1 en el espermatozoide y Cd9 en el óvulo) y la posible existencia de señales de coevolución que sugiera una interacción durante la fecundación. Para ello, se utilizó un grupo de especies de la familia Muridae cercanas filogenéticamente.

Una cuestión abierta en estudios de coevolución es qué proporción de la evolución correlacionada es atribuible a la interacción física entre dos proteínas y cual a las presiones selectivas compartidas debido a las relaciones funcionales. Un método que se puede usar para evaluar si la

coevolución observada entre dos proteínas es el resultado de cambios compensatorios para mantener una interacción consiste en analizar los dominios extracelulares esperados a estar implicados en el reconocimiento (Hakes *et al.* 2007). Aunque no se ha mostrado evidencia de una interacción física entre las proteínas Izumo1 y Cd9 (Evans 2012; Okabe 2013), los análisis realizados detectaron que la coevolución Izumo1-Cd9 es principalmente dirigida por la coevolución entre el dominio Izumo y el lazo extracelular grande de Cd9. Estos resultados sugieren que Izumo1 y Cd9 podrían interactuar directamente en *trans* durante la fusión de gametos.

En cualquier caso, se debe ser cauto a la hora de concluir que la causa potencial de una evidencia de coevolución es una interacción física. Se ha propuesto que Izumo1 y Cd9 pueden estar implicados en la organización y estabilidad de complejos de proteínas requeridos para la fusión de membranas (Ellerman *et al.* 2009; Zhu *et al.* 2002). En este escenario, si ocurren mutaciones en el lazo extracelular de Cd9, que se encarga de mediar interacciones con proteínas en la membrana del óvulo (Zhu *et al.* 2002), se podría modificar la conformación del complejo de fusión del óvulo. Las presiones selectivas podrían entonces promover cambios adaptativos en las proteínas que organizan el complejo de fusión del espermatozoide (ej: Izumo1) para compensar las modificaciones ocurridas en el complejo del óvulo, manteniendo así la afinidad entre las maquinarias de fusión de ambos gametos. Por lo tanto, podría darse un proceso de coevolución entre dos proteínas implicadas en la fusión de gametos a pesar de que no mantengan una interacción física.

Por otro lado, grupos de proteínas que participan en los mismos procesos dentro de la célula se encuentran frecuentemente sujetas a regímenes selectivos similares, lo que puede conllevar a la estimación de tasas evolutivas correlacionadas (Hakes *et al.* 2007; Pazos y Valencia 2008). La asociación funcional de Izumo1 y Cd9 por el hecho de estar implicadas en el mismo proceso de fusión espermatozoide-óvulo podría ser otra causa potencial de coevolución. Sin embargo, este no es siempre el caso, ya que para Spesp1, una proteína que supuestamente está implicada en la relocalización de Izumo1 (entre otras proteínas) (Fujihara *et al.* 2010) no mostró evidencia de coevolución con Izumo1 o Cd9.

En un modelo de interacción proteína-proteína, se asume que las dos proteínas experimentan un número similar de cambios. Sin embargo, se observó que Izumo1 mostraba una mayor tasa evolutiva que Cd9. Es posible que cambios puntuales en una proteína tengan que ser compensados por varias mutaciones en la pareja (Clark *et al.* 2009). Otra explicación podría ser que una de las proteínas interactúe con más moléculas, y por tanto su tasa evolutiva sería el producto de los cambios acumulados por las diferentes parejas. Se han identificado diversas proteínas que se asocian con Izumo1 (Sosnik *et al.* 2009; Inoue *et al.* 2010) y CD9 (Rubinstein *et al.* 2006; Ziyyat *et al.* 2006; Glazar y Evans 2009). Por lo tanto, puede que Izumo1 interactúe con un mayor número de proteínas (bien en *cis* en la membrana del espermatozoide o en *trans* con proteínas de la superficie del óvulo) que Cd9. Por otra parte, dado que Cd9 se expresa en numerosos tejidos, puede que esta

proteína tenga funciones en otras células diferentes al óvulo, y esto puede generar que Cd9 se encuentre sujeta a mayores restricciones selectivas.

La coevolución de proteínas de fecundación se ha demostrado previamente en invertebrados marinos (Clark *et al.* 2009; Hellberg *et al.* 2012). En especies de fecundación interna, donde los gametos se exponen a diversos ambientes y condiciones, y experimentan múltiples procesos fisiológicos antes de la fecundación, se podría esperar que los mecanismos de interacción de gametos tuvieran menos importancia en el mantenimiento de un sistema específico de fecundación para cada especie. No obstante, los resultados de este estudio apoyan la hipótesis de que los procesos de coevolución de proteínas que mantienen el reconocimiento gamético en las especies también ocurren en grupos con fecundación interna.

Los análisis no detectaron una correlación entre las tasas evolutivas de Izumo1 o Cd9 y los niveles de competición espermática entre las especies. Sin embargo, esto no implica que la competición espermática no esté influyendo sobre la evolución de estas proteínas. Un método alternativo para inferir las fuerzas selectivas que dirigen la evolución de proteínas reproductivas consiste en estimar la tasa relativa de divergencia de un par de proteínas que interactúan entre los gametos (Civetta 2003b; Hellberg *et al.* 2012). Este método asume que, durante la evolución de las especies más promiscuas, se esperan encontrar etapas donde las proteínas del óvulo cambian más rápidamente que las proteínas del espermatozoide debido a la selección para evitar polispermia. Mientras, en especies de baja competición espermática, el conflicto sexual es más relajado, y por tanto se espera encontrar una mayor divergencia en las proteínas del espermatozoide. Las tasas de divergencia relativa entre Izumo1 y Cd9 estimadas a través de las especies de múridos están de acuerdo con las predicciones de este modelo, sugiriendo que estas proteínas podrían encontrarse bajo la acción de fuerzas de selección sexual post-cópula.

Numerosos estudios han analizado el efecto de las fuerzas de selección sexual sobre la divergencia de las proteínas reproductivas a nivel de secuencia codificante. Sin embargo, varios trabajos han propuesto que las fuerzas selectivas pueden promover cambios en la regulación de la expresión génica (Meiklejohn *et al.* 2003; Khaitovich *et al.* 2005; Voolstra *et al.* 2007; Martín-Coello *et al.* 2009). En este contexto, la proteómica constituye una herramienta ideal para analizar los cambios en expresión que ocurren en las células, particularmente en el espermatozoide, que carece de actividad transcripcional y traduccional (Oliva *et al.* 2009). En un último trabajo se llevó a cabo un estudio de proteómica cuantitativa en el que se analizó y comparó el proteoma de tres especies próximas de roedores con diferentes niveles de competición espermática. Para ello se emplearon dos técnicas diferentes: una basada en el uso de geles en 2D (2D-DIGE) y otra basada en el recuento de péptidos usando espectrometría de masas (LC-MS/MS). Aunque los métodos basados en LC-MS/MS

ofrecen mayor resolución que aquellos basados en geles 2D, ambos pueden identificar grupos de proteínas complementarias (de Mateo *et al.* 2011).

Nuestros análisis detectaron diferencias cuantitativas entre una especie de baja competición espermática (*Mus musculus*) y dos especies de alta competición espermática (*Mus spretus* y *Mus spicilegus*) en un grupo de proteínas importantes para la función del espermatozoide y el éxito reproductivo. Así, el espermatozoide de *M. musculus* presenta una mayor abundancia de proteínas que componen el flagelo. Esta observación no se puede explicar fácilmente. Quizás se pueda asociar a que la especie *M. musculus* presente un flagelo con propiedades estructurales o funcionales diferentes a las de los espermatozoides de las otras dos especies. Estudios futuros sobre la organización estructural del flagelo y la regulación del movimiento espermático en estas especies podrían dar luz sobre esta cuestión.

Por otro lado, *M. spretus* y *M. spicilegus* mostraron una mayor abundancia para un conjunto de proteínas implicadas en la interacción espermatozoide-óvulo. Teniendo en cuenta que numerosos rasgos reproductivos que incrementan el éxito de fecundación evolucionan bajo la influencia de las fuerzas de selección sexual (Pizarri y Parker 2009), se asume que la competición espermática también selecciona aquellos espermatozoides que mejor se unen e interaccionan con los componentes de la superficie del óvulo (Swanson y Vacquier 2002). De acuerdo con esta idea, se han identificado un elevado número de proteínas implicadas en la interacción espermatozoide-óvulo que evolucionan bajo selección positiva (Swanson y Vacquier 2002; Dorus *et al.* 2010), aunque aún no se ha mostrado evidencia empírica de que cambios adaptativos en estas proteínas se asocian con una mayor eficacia del espermatozoide en unirse e interaccionar con los componentes de la membrana del óvulo. En este contexto, un posible mecanismo para incrementar la eficacia de la fecundación en especies sujetas a altos niveles de competición espermática podría ser a través de incrementar la expresión de receptores del espermatozoide para la zona pelúcida. De este modo, una mayor concentración de estos receptores en la superficie del espermatozoide podría potenciar la interacción entre los gametos, conllevando a una fecundación más rápida y eficiente.

Algunas proteínas implicadas en procesos como la motilidad espermática o la reacción acrosómica también mostraron una mayor abundancia en las especies más promiscuas. Teniendo en cuenta el importante papel de estos procesos en la competición espermática, los resultados apoyan la hipótesis de que las fuerzas de selección sexual pueden favorecer la mayor expresión de proteínas del espermatozoide con fines adaptativos. Por otro lado, los resultados sugieren que la competición espermática también puede favorecer una reducción de la expresión de las proteínas reproductivas en algunos casos.

De acuerdo con lo comentado anteriormente, los análisis basados en LC-MS/MS identificaron un número proporcionalmente mayor de proteínas que el identificado usando geles 2D. No obstante,

a partir de los geles 2D se detectaron proteínas que no fueron identificadas por LC-MS/MS y viceversa. Aquellas proteínas que revelaron diferencias significativas de abundancia en los dos métodos mostraron resultados concordantes en la mayoría de los casos, lo que apoya la robustez de los resultados. Estas observaciones nos llevan a aconsejar el uso conjunto de las dos metodologías para estudios futuros.

Con esta serie de trabajos se ha apoyado la idea de que existen numerosas proteínas expresadas en tejidos reproductivos (particularmente en el espermatozoide) que evolucionan con gran rapidez y dirigidas por selección positiva. Entre los diferentes mecanismos que se han propuesto para explicar la rápida evolución de las proteínas reproductivas, las fuerzas de selección sexual post-cópula (concretamente la competición espermática en el caso de proteínas masculinas), a menudo se han considerado como causas potenciales. No obstante, pocos estudios han sido capaces de revelar evidencia de un efecto directo de la selección sexual post-cópula sobre la evolución de las proteínas reproductivas. Esto no implica que la selección sexual sea una fuerza poco influyente en la evolución de las proteínas reproductivas, sino que el fallo en detectar el efecto de la selección sexual puede deberse a varias razones.

Por una parte, en las proteínas que muestran evidencia de selección positiva, existe gran dificultad para distinguir entre la porción de selección que se debe a la selección sexual y la que es causada por otras fuerzas selectivas (Wong 2011). Por otro lado, la no detección del efecto de la selección sexual sobre la evolución de proteínas reproductivas puede atribuirse a cuestiones metodológicas. La mayoría de los estudios que han intentado establecer asociaciones entre selección sexual y evolución de proteínas se han basado en comparaciones de tasas de evolución molecular entre especies con diferentes sistemas de apareamiento. Sin embargo, estos métodos pueden presentar algunas limitaciones. Uno de los problemas se debe a que los datos evolutivos y los datos fenotípicos que se comparan no suelen tener la misma escala, ya que estos últimos normalmente son medidas de caracteres realizadas en el presente (ej: masa relativa de testículo), mientras que las estimas de ω representan el valor promedio de una o varias ramas de un árbol filogenético (Ramm *et al.* 2008). Aunque en los últimos años se han desarrollado modelos capaces de comparar simultáneamente evolución molecular y evolución de un carácter fenotípico (O'Connor y Mundy 2009; Lartillot y Poujol 2011; Mayrose y Otto 2011), estos métodos presentan a su vez ciertas limitaciones. Además, el análisis de proteínas reproductivas usando estos métodos no ha revelado resultados significativamente diferentes respecto a estudios previos (O'Connor y Mundy 2009).

Otro problema de la mayoría de los estudios que analizan la relación entre niveles de selección sexual y la evolución de proteínas reproductivas es que normalmente se analiza un gen o unos pocos genes. Por tanto, en estos análisis se asume que las fuerzas selectivas actúan sobre el mismo gen en todas las especies analizadas. No obstante, puede ocurrir que la selección sexual actúe sobre

diferentes genes en las diferentes especies. Esto se debe tener especialmente en cuenta cuando se pretende asociar la evolución de genes con la evolución de caracteres reproductivos, ya que estos rasgos normalmente son controlados por múltiples genes y por lo tanto pueden existir múltiples proteínas modulando su adaptación (Ramm *et al.* 2008; Wong 2011). Las alternativas para evitar este problema consistirían en comparar la evolución promedio de muchos genes (en el caso de rasgos reproductivos, el máximo número de genes que se conozcan participando en tal carácter) entre especies con diferentes niveles de selección sexual, o bien desarrollar estudios de evolución molecular a nivel intraespecífico (Markow y O'Grady 2005; Findlay y Swanson 2010; Wong 2010).

Otra cuestión importante en relación a estos análisis es que principalmente se asume que la tasa de divergencia y el nivel de selección positiva en las proteínas reproductivas van a aumentar con la intensidad de selección sexual. No obstante, la selección sexual puede actuar en otras direcciones, como por ejemplo evitando la degeneración de proteínas bajo relajación funcional que pueden aportar una función importante, aunque no esencial, en la reproducción. En este caso, cabe esperar una correlación negativa entre la tasa evolutiva de la proteína y los niveles de selección sexual. Un ejemplo es la evidencia encontrada para la protamina 2 (Prm2) en roedores, que es una proteína que presenta una función redundante con su paróloga protamina 1 (Prm1) y por tanto se encuentra bajo relajación funcional, llegando a perder su expresión en algunas especies (Corzett *et al.* 1999, 2002). Sin embargo, se ha demostrado que en las especies con mayores niveles de competición espermática, Prm2 se encuentra más conservada, manteniendo la funcionalidad de la proteína (Lüke *et al.* 2011).

La mayoría de los estudios de evolución de proteínas reproductivas se han centrado en analizar la evolución de las secuencias codificantes de los genes. Sin embargo, se ha mostrado evidencia de que la rápida evolución y la selección positiva también afecta a la expresión de genes reproductivos (Khaitovich *et al.* 2005; Voolstra *et al.* 2007). De hecho, un estudio demostró que la divergencia del promotor de *Prm2* se incrementa con los niveles de competición espermática en muridos (Martín-Coello *et al.* 2009). El estudio aquí presentado sobre proteómica del espermatozoide es un claro ejemplo de que la selección sexual es capaz de promover cambios en la expresión de proteínas reproductivas con fines adaptativos. Del mismo modo, se ha prestado poca atención a la evolución estructural de las proteínas reproductivas, lo cual, como se ha demostrado en uno de los trabajos, puede ser un mecanismo de acción importante para las fuerzas de selección sexual.

En resumen, los trabajos presentados en esta memoria han servido para mostrar evidencias de la gran importancia que tienen las fuerzas de selección sexual en la evolución de las proteínas reproductivas. En el futuro, la aplicación de nuevos métodos analíticos y el diseño de estudios comparativos que tengan en cuenta el efecto heterogéneo de las fuerzas selectivas nos permitirá examinar con mayor robustez las causas de la rápida evolución de las proteínas reproductivas, así como los mecanismos de adaptación de los rasgos reproductivos.

6. CONCLUSIONES

CONCLUSIONES EN ESPAÑOL

1. Las proteínas integrales del espermatozoide de ratón muestran diferencias en la tasa evolutiva y en función de los diferentes procesos que experimenta el espermatozoide desde la espermatogénesis hasta la fecundación.
2. Una alta proporción de proteínas implicadas en la interacción de gametos y la motilidad espermática presentan evolución adaptativa, sugiriendo que estos dos procesos son focos de acción importantes de las fuerzas de selección sexual.
3. La competición espermática favorece un acortamiento de la región amino-terminal de la proteína Catsper1 en roedores. Tal acortamiento está asociado a incrementos en la velocidad de natación del espermatozoide.
4. La tasa evolutiva de PKDREJ, una proteína implicada en regulación del transporte y la reacción acrosómica del espermatozoide, incrementa con los niveles de competición espermática en roedores.
5. Dos proteínas esenciales en la fusión de gametos (Izumo1 y Cd9) muestran una correlación de sus tasas evolutivas en roedores, dando evidencia de un proceso de coevolución entre ellas.
6. La selección positiva detectada en los dominios extracelulares de Izumo1 y Cd9 y la coevolución entre estos sugiere una unión directa durante la fusión de gametos, aunque la señal de coevolución podría ser resultado de la asociación funcional entre estas proteínas.
7. Análisis de proteómica cuantitativa entre especies de roedores con diferentes niveles de competición espermática revelaron cambios en abundancia para algunas proteínas del espermatozoide.
8. Proteínas con mayor abundancia en las especies de alta competición espermática se asocian a funciones importantes para el espermatozoide el éxito reproductivo.

CONCLUSIONS IN ENGLISH

1. Integral mouse sperm proteins show differences in their evolutionary rate and levels of positive selection according to the different processes undergone by spermatozoa, from spermatogenesis to fertilization.
2. A high proportion of proteins involved in sperm-egg interaction and sperm motility evolve under positive selection. These processes may be important targets for sexual selection.
3. Sperm competition favors a shortening of the amino-terminal region of Catsper1 protein in rodents. Such shortening is associated with increases in sperm swimming velocity.
4. The evolutionary rate of PKDREJ, a protein involved in the regulation of sperm transport and acrosome reaction, increases with levels of sperm competition in rodents.
5. Two essential proteins in gamete fusion (Izumo1 and Cd9) show correlated evolutionary rates in rodents; this allows us to conclude that these proteins undergo coevolutionary processes.
6. The signal of positive selection detected in their extracellular domains, and their coevolution, suggest a possible direct binding during gamete fusion, although the signal of coevolution may be result of a functional association between these proteins.
7. Analyses of quantitative proteomics between species of rodents exhibiting different levels of sperm competition have shown significant changes in abundance of some sperm proteins.
8. The proteins with higher abundance in the species with high sperm competition levels are associated with important sperm functions) and reproductive success.

BIBLIOGRAFÍA

- Anderson, M. J., A. F. Dixson. 2002. Motility and the midpiece in primates. *Nature* 416:496.
- Anderson, M. J., J. Nyholt, A. F. Dixson. 2005. Sperm competition and the evolution of sperm midpiece volume in mammals. *J. Zool.* 267:135–42.
- Anisimova, M., J. P. Bielawski, Z. Yang. 2001. Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. *Mol. Biol. Evol.* 18:1585–92.
- Armbrust, E. V., H. M. Galindo. 2001. Rapid Evolution of a Sexual Reproduction Gene in Centric Diatoms of the Genus *Thalassiosira*. *Appl. Environ. Microbiol.* 67:3501–13.
- Avella, M. A., B. Xiong, J. Dean. 2013. The molecular basis of gamete recognition in mice and humans. *Mol. Hum. Reprod.* 19:279–89.
- Baker, M. A., G. Reeves, L. Hetherington, J. Müller, I. Baur, R. J. Aitken. 2007. Identification of gene products present in Triton X-100 soluble and insoluble fractions of human spermatozoa lysates using LC-MS/MS analysis. *Proteomics. Clin. Appl.* 1:524–32.
- Baker, M. A., L. Hetherington, G. M. Reeves, R. J. Aitken. 2008a. The mouse sperm proteome characterized via IPG strip prefractionation and LC-MS / MS identification. *Proteomics* 8:1720–30.
- Baker, M. A., L. Hetherington, G. Reeves, J. Müller, R. J. Aitken. 2008b. The rat sperm proteome characterized via IPG strip prefractionation and LC-MS/MS identification. *Proteomics* 8:2312–21.
- Balhorn, R. 2007. The protamine family of sperm nuclear proteins. *Genome. Biol.* 8.
- Beck, J. A., S. Lloyd, M. Hafezparast, M. Lennon-Pierce, J. T. Eppig, M. F. Festing, E. M. Fisher. 2000. Genealogies of mouse inbred strains. *Nat. Genet.* 24:23–5.
- Berlin, S., L. Qu, H. Ellegren. 2008. Adaptive evolution of gamete-recognition proteins in birds. *J. Mol. Evol.* 67:488–96.
- Bielawski, J. P., Z. Yang. 2001. Positive and negative selection in the DAZ gene family. *Mol. Biol. Evol.* 18:5232–9.
- Birkhead, T. R., J. G. Mart, T. Burke, D. P. Froman. 1999. Sperm mobility determines the outcome of sperm competition in the domestic fowl. *Proc. Biol. Sci.* 266:1759–64.
- Birkhead, T. R., A. P. Moller. 1998. *Sperm Competition and Sexual Selection*. Academic P. San Diego, CA.
- Birkhead, T. R., A. P. Moller. 1992. *Sperm Competition in Birds*. Academic p. London.
- Bouyain, S., P. A. Longo, S. Li, K. M. Ferguson, D. J. Leahy. 2005. The extracellular region of ErbB4 adopts a tethered conformation in the absence of ligand. *Proc. Natl. Acad. Sci. USA.* 102:15024–29.
- Braisted, J. C., S. Kuntumalla, C. Vogel, E. M. Marcotte, A. R. Rodrigues, R. Wang, S.-T. Huang, E. S. Ferlanti, A. I. Saeed, R. D. Fleischmann, S. N. Peterson, R. Pieper. 2008. The APEX

Quantitative Proteomics Tool: generating protein quantitation estimates from LC-MS/MS proteomics results. *BMC Bioinformatics* 9:529.

Breed, W. G., J. Taylor. 2000. Body mass, testes mass and sperm size in murine rodents. *J. Mammal.* 81:758–68.

Briskie, J. V., M. Robert. 1992. Sperm size and sperm competition in birds. *Proc. Biol. Sci.* 247:89–95.

Byrne, P. G., L. W. Simmons, J. D. Roberts. 2003. Sperm competition and the evolution of gamete morphology in frogs. *Proc. R. Soc. Lond.* 270:2079–86.

Cai, X., D. E. Clapham. 2008. Evolutionary Genomics Reveals Lineage-Specific Gene Loss and Rapid Evolution of a Sperm-Specific Ion Channel Complex : CatSper and CatSperb. *PLoS One* 3:e3569.

Cameron, E., T. Day, L. Rowe. 2007. Sperm competition and the evolution of ejaculate composition. *Am. Nat.* 169:E158–72.

Cao, W., G. L. Gerton, S. B. Moss. 2006. Proteomic Profiling of Accessory Structures from the Mouse Sperm Flagellum. *Mol. Cell. Proteomics* 5:801–10.

Carlson, A. E., T. A. Quill, R. E. Westenbroek, S. M. Schuh, B. Hille, D. F. Babcock. 2005. Identical Phenotypes of CatSper1 and CatSper2 Null Sperm. *J. Biol. Chem.* 280:32238–44.

Carlson, A. E., R. E. Westenbroek, T. Quill, D. Ren, D. E. Clapham, B. Hille, D. L. Garbers, D. F. Babcock. 2003. CatSper1 required for evoked Ca²⁺ entry and control of flagellar function in sperm. *P. Natl. Acad. Sci. USA.* 100:1–5.

Casselton, L. A., N. S. Olesnick. 1998. Molecular genetics of mating recognition in basidiomycete fungi. *Microbiol. Mol. Biol. Rev.* 62:55–70.

Chanchevalap, S. 2000. Involvement of Histidine Residues in Proton Sensing of ROMK1 Channel. *J. Biol. Chem.* 275:7811–17.

Chapman, T. 2001. Seminal fluid-mediated fitness traits in *Drosophila*. *Heredity (Edinb).* 87:511–21.

Chapman, T., J. Bangham, L. Rowe. 2003. Sexual conflict. *Evolution (N. Y).* 18:4–6.

Civetta, A. 2003a. Positive Selection Within Sperm-Egg Adhesion Domains of Fertilin: An ADAM Gene with a Potential Role in Fertilization. *Mol. Biol. Evol.* 20:21–9.

Civetta, A. 2003b. Shall we dance or shall we fight? Using DNA sequence data to untangle controversies surrounding sexual selection 1 , 2. *Genome* 46:925–29.

Civetta, A., R. S. Singh. 1995. High Divergence of Reproductive Tract Proteins and Their Association with Postzygotic Reproductive Isolation in *Drosophila melanogaster* and *Drosophila virilis* Group Species. *J. Mol. Evol.* 41:1085–95.

Clark, G. F., A. Dell. 2006. Molecular Models for Murine. *J. Biol. Chem.* 281:13853–56.

- Clark, N. L., J. E. Aagaard, W. J. Swanson. 2006. Evolution of reproductive proteins from animals and plants. *Reproduction* 131:11–22.
- Clark, N. L., J. Gasper, M. Sekino, S. A. Springer, C. F. Aquadro, W. J. Swanson. 2009. Coevolution of interacting fertilization proteins. *PLoS Genet.* 5:e1000570.
- Clark, N. L., W. J. Swanson. 2005. Pervasive adaptive evolution in primate seminal proteins. *PLoS Genet.* 1:e35.
- Corzett, M., C. Kramer, R. Blacher, J. O. E. Mazrimas, R. O. D. Balhorn. 1999. Analysis of Hamster Protamines : Primary. *Mol. Reprod. Dev.* 282:273–82.
- Corzett, M., J. O. E. Mazrimas, R. O. D. Balhorn. 2002. Protamine 1 : Protamine 2 Stoichiometry in the Sperm of Eutherian Mammals. *Mol. Reprod. Dev.* 527.
- Coulter, K. L., F. Perier, C. M. Redeke, C. A. Vanderberg. 1995. Identification and molecular localization of a pH-sensing domain for the inward rectifier potassium channel HIR. *Neuron.* 15:1157–68.
- Craig, R., R. C. Beavis. 2004. TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* 20:1466–67.
- Darwin, C. 1871. *The Descent of Man, selection in relation to sex.* Albemarle. London.
- De Jong, W. W., L. Rydén. 1981. Causes of more frequent deletions than insertions in mutations and protein evolution. *Nature* 290:157–59.
- De Mateo, S., J. Castillo, J. M. Estanyol, J. L. Ballescà, R. Oliva. 2011. Proteomic characterization of the human sperm nucleus. *Proteomics* 11:2714–26.
- Dean, M. D., N. L. Clark, G. D. Findlay, R. C. Karn, X. Yi, W. J. Swanson, M. J. MacCoss, M. W. Nachman. 2009. Proteomics and comparative genomic investigations reveal heterogeneity in evolutionary rate of male reproductive proteins in mice (*Mus domesticus*). *Mol. Biol. Evol.* 26:1733–43.
- Dean, M. D., J. M. Good, M. W. Nachman. 2008. Adaptive evolution of proteins secreted during sperm maturation: an analysis of the mouse epididymal transcriptome. *Mol Biol Evol.* 25:383–92.
- DelBarco-Trillo, J., M. H. Ferkin. 2004. Male mammals respond to a risk of sperm competition conveyed by odours of conspecific males. *Nature* 431:446–49.
- Delmas, P. 2005. Polycystins: polymodal receptor/ion-channel cellular sensors. *Pflugers Arch.* 451:264–76.
- Dewsbury, D. A. 1984. Sperm competition in muroid rodents. Pp. 547–571 in R. L. Smith, ed. *Sperm competition and the evolution of animal mating systems.* New York.
- Dorus, S., S. A. Busby, U. Gericke, J. Shabanowitz, D. F. Hunt, T. L. Karr. 2006. Genomic and functional evolution of the *Drosophila melanogaster* sperm proteome. *Nat. Genet.* 38:1440–45.
- Dorus, S., P. D. Evans, G. J. Wyckoff, S. S. Choi, B. T. Lahn. 2004. Rate of molecular evolution of the seminal protein gene SEMG2 correlates with levels of female promiscuity. *Nat. Genet.* 36:1326–29.

- Dorus, S., S. Skerget, T. L. Karr. 2012. Proteomic discovery of diverse immunity molecules in mammalian spermatozoa. *Sys. Biol. Repr. Med.* 58:218–28.
- Dorus, S., E. R. Wasbrough, J. Busby, E. C. Wilkin, T. L. Karr. 2010. Sperm proteomics reveals intensified selection on mouse sperm membrane and acrosome genes. *Mol. Biol. Evol.* 27:1235–46.
- Eberhard, W. G. 1996. *Female Control: Sexual Selection by Cryptic Female Choice*. Princeton. Princeton.
- Eden, E., R. Navon, I. Steinfeld, D. Lipson, Z. Yakhini. 2009. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* 10:48.
- Escalier, D. 2006. Knockout mouse models of sperm flagellum anomalies. *Hum. Reprod.* 12:449–61.
- Escalier, D., A. Touré. 2012. Malformations L1 l'appareil de flagellaire du spermatozoïde impliquées dans l'infertilité chez l'homme. *Médecine/ciences* 28:503–11.
- Evans, J. P. 2012. Sperm-egg interaction. *Annu. Rev. Physiol.* 74:477–502.
- Evans, J. P. 2002. The molecular basis of sperm-oocyte membrane interactions during mammalian fertilization. *Hum. Reprod. Update* 8:297–311.
- Fabre, P.-H., L. Hautier, D. Dimitrov, E. J. Douzery. 2012. A glimpse on the pattern of rodent diversification: a phylogenetic approach. *BMC Evol. Biol.* 12:88.
- Ferris, P. J., C. Pavlovic, S. Fabry, U. W. Goodenough. 1997. Rapid evolution of sex-related genes in *Chlamydomonas*. *Proc. Natl. Acad. Sci. USA.* 94:8634–39.
- Findlay, G. D., W. J. Swanson. 2010. Proteomics enhances evolutionary and functional analysis of reproductive proteins. *BioEssays.* 32:26–36.
- Finn, S., A. Civetta. 2010. Sexual selection and the molecular evolution of ADAM proteins. *J. Mol. Evol.* 71:231–40.
- Fitzpatrick, J. L., R. Montgomerie, J. K. Desjardins, K. A. Stiver, N. Kolm, S. Balshine. 2009. Female promiscuity promotes the evolution of faster sperm in cichlid fishes. *P. Natl. Acad. Sci. USA.* 106:1128–32.
- Fujihara, Y., M. Murakami, N. Inoue, Y. Satouh, K. Kaseda, M. Ikawa, M. Okabe. 2010. Sperm equatorial segment protein 1, SPESP1, is required for fully fertile sperm in mouse. *J. Cell. Sci.* 123:1531–36.
- Gage, M. J. G. 1994. Associations between Body Size, Mating Pattern, Testis Size and Sperm Lengths across Butterflies. *Proc. R. Soc. B Biol. Sci.* 258:247–54.
- Gage, M. J. G., C. P. Macfarlane, S. Yeates, R. G. Ward, J. B. Searle, G. A. Parker. 2004. Spermatozoal Traits and Sperm Competition in Atlantic Salmon. *Curr Biol.* 14:44–7.
- Galewski, T., M. Tilak, S. Sanchez, P. Chevret, E. Paradis, E. J. P. Douzery. 2006. The evolutionary radiation of Arvicolinae rodents (voles and lemmings): relative contribution of nuclear and mitochondrial DNA phylogenies. *BMC Evol. Biol.* 17:1–17.

- Galindo, B. E., V. D. Vacquier, W. J. Swanson. 2003. Positive selection in the egg receptor for abalone sperm lysin. *Proc. Natl. Acad. Sci. USA*. 100:4639–43.
- García-Gonzalez, F., L. W. Simmons. 2005. Sperm Viability Matters in Insect Sperm Competition. *Curr. Biol.* 15:271–75.
- García-González, F., L. W. Simmons. 2007. Shorter sperm confer higher competitive fertilization success. *Evolution*. 61:816–24.
- Gasparini, C., L. W. Simmons, M. Beveridge, J. P. Evans. 2010. Sperm Swimming Velocity Predicts Competitive Fertilization Success in the Green Swordtail *Xiphophorus helleri*. *PLoS ONE* 5:e12146.
- Gasper, J., W. J. Swanson. 2006. Molecular population genetics of the gene encoding the human fertilization protein zonadhesin reveals rapid adaptive evolution. *Am. J. Hum. Genet.* 79:820–30.
- Gavrilets, S. 2000. Rapid evolution of reproductive barriers driven by sexual conflict. *Nature* 403:886–89.
- Glabe, C. G., W. J. Lennarz. 1979. Species-specific sperm adhesion in sea urchins. A quantitative investigation of bindin-mediated egg agglutination. *J. Cell. Biol.* 83:595–604.
- Glassey, B., A. Civetta. 2004. Positive Selection at Reproductive ADAM Genes with Potential Intercellular Binding Activity. *Mol. Biol. Evol.* 21:851–59.
- Glazar, A., J. P. Evans. 2009. Immunoglobulin superfamily member IgSF8 (EWI-2) and CD9 in fertilisation: evidence of distinct functions for CD9 and a CD9-associated protein in mammalian sperm – egg interaction. *Repr. Fert. Dev.* 21:293–303.
- Goios, A., L. Pereira, M. Bogue, V. Macaulay, A. Amorim. 2007. mtDNA phylogeny and evolution of laboratory mouse strains. *Genome. Res.* 17:293–98.
- Goldman, N., Z. Yang. 1994. A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Mol. Biol. Evol.* 11:725–36.
- Gomendio, M., A. H. Harcourt, E. R. S. Roldan. 1998. Sperm competition in mammals. Pp. 667–751 in T. R. Birkhead and A. P. Moller, eds. *Sperm Competition and Sexual Selection*. London.
- Gomendio, M., J. Martin-coello, C. Crespo, C. Magaña, E. R. S. Roldan. 2006. Sperm competition enhances functional capacity of mammalian spermatozoa. *P. Natl. Acad. Sci. USA*. 103:15113–17.
- Gomendio, M., E. R. S. Roldan. 2008. Implications of diversity in sperm size and function for sperm competition and fertility. *Int. J. Dev. Biol.* 52:439–47.
- Gómez Montoto, L., C. Magaña, M. Tourmente, J. Martín-Coello, C. Crespo, J. J. Luque-Larena, M. Gomendio, E. R. S. Roldan. 2011a. Sperm competition, sperm numbers and sperm quality in muroid rodents. *PLoS one*. 6:e18173.
- Gómez Montoto, L., M. Varea Sánchez, M. Tourmente, J. Martín-Coello, J. J. Luque-Larena, M. Gomendio, E. R. S. Roldan. 2011b. Sperm competition differentially affects swimming velocity and size of spermatozoa from closely related muroid rodents: head first. *Reproduction*. 142:819–30.

- Good, J. M., M. W. Nachman. 2005. Rates of protein evolution are positively correlated with developmental timing of expression during mouse spermatogenesis. *Mol. Biol. Evol.* 22:1044–52.
- Goodman, M., C. A. Porter, J. Czelusniak, S. L. Page, H. Schneider, J. Shoshani, G. Gunnell, C. P. Groves. 1998. Toward a phylogenetic classification of Primates based on DNA evidence complemented by fossil evidence. *Mol. Phylogenet. Evol.* 9:585–98.
- Goudet, G., S. Mugnier, I. Callebaut, P. Monget. 2008. Phylogenetic Analysis and Identification of Pseudogenes Reveal a Progressive Loss of Zona Pellucida Genes During Evolution of Vertebrates. *J. Biol. Reprod.* 78:796–806.
- Guindon, S., J.-F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, O. Gascuel. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59:307–21.
- Hakes, L., S. C. Lovell, S. G. Oliver, D. L. Robertson. 2007. Specificity in protein interactions and its relationship with sequence diversity and coevolution. *P. Natl. Acad. Sci. USA* 104:7999–8004.
- Hamm, D., B. S. Mautz, M. F. Wolfner, C. F. Aquadro, W. J. Swanson. 2007. Evidence of amino acid diversity-enhancing selection within humans and among primates at the candidate sperm-receptor gene PKDREJ. *Am. J. Hum. Genet.* 81:44–52.
- Harris, C. L., M. Mizuno, B. P. Morgan. 2006. Complement and complement regulators in the male reproductive system. *Mol. Immunol.* 43:57–67.
- Hellberg, M. E., A. B. Dennis, P. Arbour-Reily, J. E. Aagaard, W. J. Swanson. 2012. The Tegula tango: a coevolutionary dance of interacting, positively selected sperm and egg proteins. *Evolution* 66:1681–94.
- Herlyn, H., H. Zischler. 2005. Identification of a positively evolving putative binding region with increased variability in posttranslational motifs in zonadhesin MAM domain 2. *Mol. Phyl. Evol.* 37:62–72.
- Herlyn, H., H. Zischler. 2007. Sequence evolution of the sperm ligand zonadhesin correlates negatively with body weight dimorphism in primates. *Evolution* 61:289–98.
- Herlyn, H., H. Zischler. 2008. The molecular evolution of sperm zonadhesin. *Int. J.* 790:781–90.
- Ho, K., C. A. Wolff, S. S. Suarez. 2009. CatSper-null mutant spermatozoa are unable to ascend beyond the oviductal reservoir. *Repr. Fert. Dev.* 21:345–50.
- Hodgson, D. J. Å., D. J. Hosken. 2006. Sperm competition promotes the exploitation of rival ejaculates. *Ecology* 243:230–34.
- Honda, A., J. Siruntawinetti, T. Baba. 2002. Role of acrosomal matrix proteases in sperm → zona pellucida interactions. *Hum. Reprod.* 8:405–12.
- Hong, S.-J., P. C.-N. Chiu, K.-F. Lee, J. Y.-M. Tse, P.-C. Ho, W. S.-B. Yeung. 2009. Cumulus cells and their extracellular matrix affect the quality of the spermatozoa penetrating the cumulus mass. *Fertil. Steril.* 92:971–78.

- Hoshi, T., W. N. Zagotta, R. W. Aldrich. 1990. Biophysical and Molecular Mechanisms of Shaker Potassium Channel Inactivation. *Science*. 250:533–38.
- Huang, Y. H., Y. H. Chen, C. M. Lin, Y. Y. Ciou, S. P. Kuo, C. T. Chen, C. M. Shih, E. E. Chang. 2007. Suppression effect of seminal vesicle autoantigen on platelet-activating factor-induced mouse sperm capacitation. *J. Cell. Biochem*. 100:941–51.
- Huerta-Cepas, J., J. Dopazo, T. Gabaldón. 2010. ETE: a python Environment for Tree Exploration. *BMC Bioinformatics* 11:24.
- Humphries, S., J. P. Evans, L. W. Simmons. 2008. Sperm competition: linking form to function. *BMC Evol. Biol.* 8:319.
- Hunter, F. M., T. R. Birkhead. 2002. Sperm viability and sperm competition in insects. *Curr. Biol.* 12:121–23.
- Ikawa, M., N. Inoue, M. Okabe. 2008. Mechanisms of sperm-egg interactions emerging from gene-manipulated animals. *Int. J. Dev. Biol.* 52:657–64.
- Immler, S., M. Saint-Jalme, L. Lesobre, G. Sorci, Y. Roman, T. R. Birkhead. 2007. The evolution of sperm morphometry in pheasants. *J. Evol. Biol.* 20:1008–14.
- Inoue, N., M. Ikawa, A. Isotani, M. Okabe. 2005. The immunoglobulin superfamily protein Izumo is required for sperm to fuse with eggs. *Nature* 434:4–8.
- Inoue, N., T. Kasahara, M. Ikawa, M. Okabe. 2010. Identification and disruption of sperm-specific angiotensin converting enzyme-3 (ACE3) in mouse. *PLoS One* 5:e10301.
- Jaiswal, B. S., M. Eisenbach. 2002. Capacitation. Pp. 57–117 in D. M. Hardy, ed. *Fertilization*. San Diego, CA.
- Jansa, S. A., B. L. Lundrigan, P. K. Tucker. 2003. Tests for positive selection on immune and reproductive genes in closely related species of the murine genus *mus*. *J. Mol. Evol.* 56:294–307.
- Jennions, M. D., N. I. Passmore. 1993. Sperm competition in frogs: testis size and a sterile male experiment on *Chiromantis xerampelina*.pdf. *Biol. J. Linn. Soc.* 50:211-20.
- Jensen-Seaman, M. I., W.-H. Li. 2003. Evolution of the hominoid semenogelin genes, the major proteins of ejaculated semen. *J. Mol. Evol.* 57:261–70.
- Jin, M., E. Fujiwara, Y. Kakiuchi, M. Okabe, Y. Satouh, S. A. Baba. 2011. Most fertilizing mouse spermatozoa begin their acrosome reaction before contact with the zona pellucida during in vitro fertilization. *PNAS* 108:4892–96.
- Kachroo, A., C. R. Schopfer, M. E. Nasrallah. 2001. Allele-Specific Receptor-Ligand Interactions in Brassica Self-Incompatibility. *Science*. 293:1824–26.
- Kamei, N., C. G. Glabe. 2003. The species-specific egg receptor for sea urchin sperm adhesion is EBR1, a novel ADAMTS protein. *Genes Dev.* 17:2502–07.
- Karp, N. A., K. S. Lilley. 2005. Maximising sensitivity for detecting changes in protein expression: experimental design using minimal CyDyes. *Proteomics* 5:3105–15.

- Karp, N. A., P. S. McCormick, M. R. Russell, K. S. Lilley. 2007. Experimental and statistical considerations to avoid false conclusions in proteomics studies using differential in-gel electrophoresis. *Mol. Cell. Proteomics* 6:1354–64.
- Keane, T. M., L. Goodstadt, P. Danecek, M. A. White, K. Wong, B. Yalcin, A. Heger, H. Whitley, A. Agam, G. Slater, et al. 2011. Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature*. 477:5–10.
- Keller, A., J. Eng, N. Zhang, X. Li, R. Aebersold. 2005. A uniform proteomics MS/MS analysis platform utilizing open XML file formats. *Mol. Syst. Biol.* 1:2005-17.
- Kenagy, G. J., S. C. Trombulak. 1986. Size and Function of Mammalian Testes in Relation to Body Size. *J. Mammal.* 67:1–22.
- Khaitovich, P., I. Hellmann, W. Enard, K. Nowick, M. Leinweber, H. Franz, G. Weiss, M. Lachmann, S. Pääbo. 2005. Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. *Science* 309:1850–54.
- Kim, T., J. Oh, J. Woo, E. Choi, S. H. Im, Y. J. Yoo, D. H. Kim, H. Nishimura, C. Cho. 2006. Expression and Relationship of Male Reproductive ADAMs in Mouse. *Biol. Reprod.* 74:744–50.
- Kingan, S. B., M. Tatar, D. M. Rand. 2003. Reduced polymorphism in the chimpanzee semen coagulating protein, semenogelin I. *J. Mol. Evol.* 57:159–69.
- Kitadokoro, K., D. Bordo, G. Galli, R. Petracca, F. Falugi, S. Abrignani, G. Grandi, M. Bolognesi. 2001. CD81 extracellular domain 3D structure: insight into the tetraspanin superfamily structural motifs. *EMBO J.* 20:12–8.
- Kresge, N., V. D. Vacquier, C. D. Stout. 2001. Abalone lysin: the dissolving and evolving sperm protein. *Bioessays* 23:95–103.
- Krzanowska, H. 1974. The passage of abnormal spermatozoa through the uterotubal junction of the mouse. *J. Reprod. Fertil.* 38:81-90.
- Lartillot, N., R. Poujol. 2011. A phylogenetic model for investigating correlated evolution of substitution rates and continuous phenotypic characters. *Mol. Biol. Evol.* 28:729–44.
- Lee, Y. H., T. Ota, V. D. Vacquier. 1995. Positive selection is a general phenomenon in the evolution of abalone sperm lysin. *Mol. Biol. Evol.* 12:231–38.
- Leighton, D. L., C. A. Lewis. 1982. Experimental hybridization in abalones. *Int. J. Invertebr. Reprod.* 5:273–82.
- Levitan, D. R., D. L. Ferrell. 2006. Selection on Gamete Recognition Depends on Sex, Density, Genotype Frequency. *Science*. 312:267–69.
- Levitan, D. R., C. P. Terhorst, N. D. Fogarty. 2007. The risk of polyspermy in three congeneric sea urchins and its implications for gametic incompatibility and reproductive isolation. *Evolution* 61:2007–14.
- Lewis, C. A., C. F. Talbot, V. D. Vacquier. 1982. A protein from abalone sperm dissolves the egg vitelline layer by a nonenzymatic mechanism. *Dev. Biol.* 92:227–39.

- Lindberg, D. 1992. Evolution, distribution and systematics of Haliotidae. Pp. 3–18 in S. Shepherd, M. Tegner, S. Guzman del Proo, eds. *Abalone of the World: Biology, Fisheries and Culture*. Oxford.
- Litscher, E. S., Z. E. V Williams, P. M. Wassarman. 2009. Zona Pellucida Glycoprotein ZP3 and Fertilization in Mammals. *Mol. Reprod. Dev.* 76:933–41.
- Liu, J., J. Xia, K. Cho, D. E. Clapham, D. Ren. 2007. CatSper B, a Novel Transmembrane Protein in the CatSper Channel Complex. *J. Biol. Chem.* 282:18945–52.
- Lopez, L. C., E. M. Bayna, D. Litoff, N. L. Shaper, J. H. Shaper, B. D. Shur. 1985. Receptor function of mouse sperm surface galactosyltransferase during fertilization. *J. Cell Biol.* 101:1501–10.
- Lu, P., C. Vogel, R. Wang, X. Yao, E. M. Marcotte. 2007. Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nat. Biotechnol.* 25:117–24.
- Lüke, L., A. Vicens, F. Serra, J. J. Luque-Larena, H. Dopazo, E. R. S. Roldan, M. Gomendio. 2011. Sexual Selection Halts the Relaxation of Protamine 2 among Rodents. *PLoS One*. 6:e29247.
- Lüpold, S., S. Calhim, S. Immler, T. R. Birkhead. 2009. Sperm morphology and sperm velocity in passerine birds. *Proc. Biol. Sci.* 276:1175–81.
- Luporini, P., A. Vallesi, C. Miceli, R. A. Bradshaw. 1995. Chemical Signaling in Ciliates. *J. Eukaryot. Microbiol.* 42:208–12.
- Lyon, J. D., V. D. Vacquier. 1999. Regions Mediating Species-Specific Recognition of the Abalone Egg Vitelline Envelope. *Dev. Biol.* 214:151–59.
- Macek, M. B., L. C. Lopez, B. D. Shur. 1991. Aggregation of beta-1,4-galactosyltransferase on mouse sperm induces the acrosome reaction. *Dev. Biol.* 147:440–44.
- Mah, S. A., W. J. Swanson, V. D. Vacquier. 2005. Positive Selection in the Carbohydrate Recognition Domains of Sea Urchin Sperm Receptor for Egg Jelly (suREJ) Proteins. *Mol. Biol. Evol.* 22:533–41.
- Makalowski, W., M. S. Boguski. 1998. Evolutionary parameters of the transcribed mammalian genome: An analysis of 2,820 orthologous rodent and human sequences. *Proc. Natl. Acad. Sci.* 95:9407–12.
- Malo, A. F., J. Garde, B. Lang-lenton, A. J. Soler, E. R. S. Roldan. 2006. Sperm design and sperm function. *Biol. Lett.* 246–49.
- Malo, A. F., J. Julia, A. J. Soler, J. Garde, M. Gomendio, E. R. S. Roldan. 2005. Male Fertility in Natural Populations of Red Deer Is Determined by Sperm Velocity and the Proportion of Normal Spermatozoa. *Biol. Reprod.* 72:822–29.
- Markow, T. A., P. M. O’Grady. 2005. Evolutionary genetics of reproductive behavior in *Drosophila*: connecting the dots. *Annu. Rev. Genet.* 39:263–91.
- Martin-Coello, J., J. Benavent-Corai, E. R. S. Roldan, M. Gomendio. 2009. Sperm competition promotes asymmetries in reproductive barriers between closely related species. *Evolution*. 63:613–23.

- Martín-Coello, J., H. Dopazo, L. Arbiza, J. Ausió, E. R. S. Roldan, M. Gomendio. 2009. Sexual selection drives weak positive selection in protamine genes and high promoter divergence, enhancing sperm competitiveness. *Proc. R. Soc. Lond.* 276:2427–36.
- Mayrose, I., S. P. Otto. 2011. A likelihood method for detecting trait-dependent shifts in the rate of molecular evolution. *Mol. Biol. Evol.* 28:759–70.
- McCartney, M. A., H. A. Lessios. 2004. Adaptive evolution of sperm binding tracks egg incompatibility in neotropical sea urchins of the genus *Echinometra*. *Mol. Biol. Evol.* 21:732–45.
- Meiklejohn, C. D., J. Parsch, J. M. Ranz, D. L. Hartl. 2003. Rapid evolution of male-biased gene expression in *Drosophila*. *Proc. Natl. Acad. Sci. USA.* 100:9894–99.
- Metz, E. C., R. E. Kane, H. Yanagimachi, S. R. Palumbi. 1994. Fertilization Between Closely Related Sea Urchins Is Blocked by Incompatibilities During Sperm-Egg Attachment and Early Stages of Fusion. *Biol. Bull.* 187:23–34.
- Metz, E. C., S. R. Palumbi. 1996. Positive selection and sequence rearrangements generate extensive polymorphism in the gamete recognition protein binding. *Mol. Biol. Evol.* 13:397–406.
- Mi, H., P. Thomas. 2009. PANTHER pathway: an ontology-based pathway database coupled with data analysis tools. *Methods Mol. Biol.* 563:123–40.
- Miller, G. T., S. Pitnick. 2002. Sperm-female coevolution in *Drosophila*. *Science.* 298:1230–33.
- Miyado, K. 2000. Requirement of CD9 on the Egg Plasma Membrane for Fertilization. *Science.* 287:321–24.
- Moore, J. a., C. H. Daugherty, S. S. Godfrey, N. J. Nelson. 2009. Seasonal monogamy and multiple paternity in a wild population of a territorial reptile (tuatara). *Biol. J. Linn. Soc.* 98:161–70.
- Moy, G. W., L. M. Mendoza, J. R. Schulz, W. J. Swanson, C. G. Glabe, V. D. Vacquier. 1996. The sea urchin sperm receptor for egg jelly is a modular protein with extensive homology to the human polycystic kidney disease protein, PKD1. *J. Cell. Biol.* 133:809–17.
- Navarro, B., Y. Kirichok, J. Chung, D. E. Clapham. 2008. Ion channels that control fertility in mammalian spermatozoa. *Int. J. Dev. Biol.* 52:607–13.
- Neumann, K., J. Michaux, V. Lebedev, N. Yigit, E. Colak, N. Ivanova, A. Poltoraus, A. Surov, G. Markov, S. Maak et al. 2006. Molecular phylogeny of the Cricetinae subfamily based on the mitochondrial cytochrome b and 12S rRNA genes and the nuclear vWF gene. *Mol. Phylogenet. Evol.* 39:135–48.
- O'Connor, T. D., N. I. Mundy. 2009. Genotype-phenotype associations: substitution models to detect evolutionary associations between phenotypic variables and genotypic evolutionary rate. *Bioinformatics* 25:i94–100.
- Okabe, M. 2013. The cell biology of mammalian fertilization. *Development* 140:4471–79.
- Oliva, R., S. de Mateo, J. M. Estanyol. 2009. Sperm cell proteomics. *Proteomics* 9:1004–17.

- Oliva, R., G. H. Dixon. 1991. Vertebrate Protamine Genes and the Histone-to-Protamine Replacement Reaction. *Prog. Nuc. Ac. Res. Mol. Biol.* 40:25–94.
- Palumbi, S. R. 1999. All males are not created equal: fertility differences depend on gamete recognition polymorphisms in sea urchins. *Proc. Natl. Acad. Sci. USA.* 96:12632–37.
- Parker, G. A. 1970. Sperm competition and its evolutionary consequences in insects. *Biol. Rev.* 45:525–67.
- Parker, G. A. 1984. Sperm competition and the evolution of animal mating strategies. Pp. 1–60 *in* R. L. Smith, ed. *Sperm Competition and the Evolution of Animal Mating Systems*. London.
- Parker, G. A., T. Pizzari, S. P. Road. 2010. Sperm competition and ejaculate economics. *Biol. Rev.* 897–934.
- Pazos, F., A. Valencia. 2008. Protein co-evolution, co-adaptation and interactions. *EMBO J.* 27:2648–55.
- Pazos, F., A. Valencia. 2001. Similarity of phylogenetic trees as indicator of protein-protein interaction. *Protein. Eng.* 14:609–14.
- Peddinti, D., B. Nanduri, A. Kaya, J. M. Feugang, S. C. Burgess, E. Memili. 2008. Comprehensive proteomic analysis of bovine spermatozoa of varying fertility rates and identification of biomarkers associated with fertility. *BMC Syst. Biol.* 2:19.
- Pilch, B., M. Mann. 2006. Large-scale and high-confidence proteomic analysis of human seminal plasma. *Proc. R. Soc. London* 7:R40.
- Pizarri, T., G. A. Parker. 2009. Sperm competition and sperm phenotype. Pp. 207–245 *in* T. R. Birkhead, D. J. Hosken, S. Pitnick, eds. *Sperm Biology: an Evolutionary Perspective*. London.
- Podlaha, O., D. M. Webb, P. K. Tucker, J. Zhang. 2005. Positive selection for indel substitutions in the rodent sperm protein Catsper1. *Mol. Biol. Evol.* 22:1845–52.
- Podlaha, O., J. Zhang. 2003. Positive selection on protein-length in the evolution of a primate sperm ion channel. *P. Natl. Acad. Sci. USA.* 100:12241–46.
- Poiani, A. 2006. Complexity of seminal fluid: a review. *Behav. Ecol. Sociobiol.* 60:289–310.
- Primakoff, P., W. Lathrop, L. Woolman, A. Cowan, D. Myles. 1988. Fully effective contraception in male and female guinea pigs immunized with the sperm protein PH-20. *Nature* 335:543–46.
- Primakoff, P., D. G. Myles. 2000. The ADAM gene family: surface proteins with adhesion and protease. *Trends Genet.* 16:83–7.
- Qi, H., M. M. Moran, B. Navarro, J. A. Chong, G. Krapivinsky, L. Krapivinsky, Y. Kirichok, I. S. Ramsey, T. A. Quill, D. E. Clapham. 2006. All four CatSper ion channel proteins are required for male fertility and sperm cell hyperactivated motility. *P. Natl. Acad. Sci. USA.* 22:1219–23.
- Queralt, R., R. Adroer, R. Oliva, R. J. Winkfein, J. D. Retief, G. H. Dixon. 1995. Evolution of Protamine P1 Genes in Mammals. *J. Mol. Evol.* 40:601–07.

- Quill, T. A., D. L. Garbers. 2002. Sperm motility activation and chemoattraction. Pp. 29–55 in D. M. Hardy, ed. *Fertilization*. San Diego, CA.
- Ramm, S. A., L. McDonald, J. L. Hurst, R. J. Beynon, P. Stockley. 2009. Comparative proteomics reveals evidence for evolutionary diversification of rodent seminal fluid and its functional significance in sperm competition. *Mol. Biol. Evol.* 26:189–98.
- Ramm, S. A., P. L. Oliver, C. P. Ponting, P. Stockley, R. D. Emes. 2008. Sexual selection and the adaptive evolution of mammalian ejaculate proteins. *Mol. Biol. Evol.* 25:207–19.
- Ramm, S. A., G. A. Parker, P. Stockley. 2005. Sperm competition and the evolution of male reproductive anatomy in rodents. *Proc. R. Soc. London* 272:949–55.
- Ren, D., B. Navarro, G. Perez, A. C. Jackson, S. Hsu, Q. Shi, J. L. Tilly, D. E. Clapham. 2001. A sperm ion channel required for sperm motility and male fertility. *Nature* 413:603–09.
- Retief, J. D., G. H. Dixon, H. S. Centre. 1993. Evolution of pro-protamine 2 in primates. *Eur J Biochem.* 214:609–15.
- Richman, A. D., J. R. Kohn. 2000. Evolutionary genetics of self-incompatibility in the Solanaceae. *Plant Mol. Biol.* 42:169–79.
- Rodeheffer, C., B. D. Shur. 2002. Targeted mutations in β 1,4-galactosyltransferase I reveal its multiple cellular functions. *Bioch. Biophys. Acta* 1573:258–70.
- Roldan, E. R. S., M. Gomendio, A. D. Vitullo. 1992. The evolution of eutherian spermatozoa and underlying selective forces: female selection and sperm competition. *Biol. Rev.* 67:551–93.
- Ronquist, F., J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–74.
- Rooney, A. P., J. Zhang. 1999. Rapid evolution of a primate sperm protein: relaxation of functional constraint or positive Darwinian selection? *Mol. Biol. Evol.* 16:706–10.
- Rubinstein, E., A. Ziyat, M. Prenant, E. Wrobel, J. Wolf, S. Levy, C. Boucheix. 2006. Reduced fertility of female mice lacking CD81. *Dev. Biol.* 290:351–58.
- Ruggiu, M., R. Speed, M. Taggart, S. J. McKay, F. Kilanowski, P. Saunders, J. Dorin, H. J. Cooke. 1997. The mouse Dazl gene encodes a cytoplasmic protein essential for gametogenesis. *Nature* 386:73–7.
- Sandstedt, S. A., P. K. Tucker. 2005. Male-Driven Evolution in Closely Related Species of the Mouse Genus *Mus*. *J. Mol. Evol.* 5:138–44.
- Scott, M. A. 2000. A glimpse at sperm function in vivo: sperm transport and epithelial interaction in the female reproductive tract. *Anim. Reprod. Sci.* 60:337–48.
- Shackleford, T. K., N. Pound. 2006. *Sperm Competition in Humans: Classic and Contemporary Readings*. Springer-V. New York.

- Shi, Q., E. R. S. Roldan. 1995. Bicarbonate/CO₂ Is Not Required for Zona Pellucida- or Progesterone-Induced Acrosomal Exocytosis of Mouse Spermatozoa But Is Essential for Capacitation. *Biol. Reprod.* 52:540–46.
- Siegel, C., J. Castellan Jr. 1988. Non parametric statistics for the behavioural sciences. McGraw-Hill. New York.
- Simmons, L. W. 2001. Sperm competition and its evolutionary consequences in insects. Princeton . Princeton.
- Singh, R. S., R. J. Kulathinal. 2000. Sex gene pool evolution and speciation: a new paradigm. *Genes Genet. Syst.* 75:119–30.
- Skerget, S., M. Rosenow, A. Polpitiya, K. Petritis. 2013. The Rhesus Macaque (*Macaca mulatta*) Sperm Proteome. *Mol. Cell. Proteomics* 12:3052–67.
- Smith, R. L. 1984. Sperm Competition and the Evolution of Animal Mating Systems. Academic p. London.
- Sosnik, J., P. V Miranda, N. A. Spiridonov, S. Yoon, R. A. Fissore, G. R. Johnson, P. E. Visconti. 2009. Tssk6 is required for Izumo relocalization and gamete fusion in the mouse. *J. Cell Sci.* 122:2741–49.
- Stein, K. K., J. C. Go, W. S. Lane, P. Primakoff, D. G. Myles. 2006. Proteomic analysis of sperm regions that mediate sperm-egg interactions. *Proteomics* 6:3533–43.
- Steppan, S. J., R. M. Adkins, J. Anderson. 2004. Phylogeny and Divergence-Date Estimates of Rapid Radiations in Muroid Rodents. *Sys Biol.* 53:533–53.
- Steppan, S. J., R. M. Adkins, P. Q. Spinks, C. Hale. 2005. Multigene phylogeny of the Old World mice, Murinae, reveals distinct geographic lineages and the declining utility of mitochondrial genes compared to nuclear genes. *Mol. Phylogenet. Evol.* 37:370–88.
- Stockley, P., M. J. Gage, G. A. Parker, A. P. Møller. 1997. Sperm competition in fishes: the evolution of testis size and ejaculate characteristics. *Am. Nat.* 149:933–54.
- Stockley, P., B. T. Preston. 2004. Sperm competition and diversity in rodent copulatory behaviour. *Anim. Behav.* 17:1048–57.
- Suarez, S. S. 2008. Control of hyperactivation in sperm. *Hum. Reprod.* 14:647–57.
- Suarez, S. S., H.-C. Ho. 2003. Hyperactivated Motility in Sperm. *Reprod. Dom. Anim* 38:119–24.
- Sutton, K. A., M. K. Jungnickel, H. M. Florman. 2008. A polycystin-1 controls postcopulatory reproductive selection in mice. *P. Natl. Acad. Sci. USA* 105:8661–66.
- Sutton, K. A., M. K. Jungnickel, C. J. Ward, P. C. Harris, H. M. Florman. 2006. Functional Characterization of PKDREJ, a Male Germ Cell-Restricted Polycystin. *J. Cell. Physiol.* 209:493–500.
- Suzuki, H., T. Shimada, M. Terashima, K. Tsuchiya, K. Aplin. 2004. Temporal, spatial, ecological modes of evolution of Eurasian Mus based on mitochondrial and nuclear gene sequences. *Mol. Phyl. Evol.* 33:626–46.

- Swanson, W. J., A. G. Clark, H. M. Waldrip-Dail, M. F. Wolfner, C. F. Aquadro. 2001a. Evolutionary EST analysis identifies rapidly evolving male reproductive proteins in *Drosophila*. *Proc. Natl. Acad. Sci. USA*. 98:7375–79.
- Swanson, W. J., R. Nielsen, Q. Yang. 2003. Pervasive adaptive evolution in mammalian fertilization proteins. *Mol. Biol. Evol.* 20:18–20.
- Swanson, W. J., V. D. Vacquier. 1995. Extraordinary divergence and positive Darwinian selection in a fusagenic protein coating the acrosomal process of abalone spermatozoa. *Proc. Natl. Acad. Sci. USA*. 92:4957–61.
- Swanson, W. J., V. D. Vacquier. 1997. The abalone egg vitelline envelope receptor for sperm lysin is a giant multivalent molecule. *Proc. Natl. Acad. Sci. USA*. 94:6724–29.
- Swanson, W. J., V. D. Vacquier. 2002. The rapid evolution of reproductive proteins. *Genetics*. 3:137–44.
- Swanson, W. J., A. Wong, M. F. Wolfner, C. F. Aquadro. 2004. Evolutionary expressed sequence tag analysis of *Drosophila* female reproductive tracts identifies genes subjected to positive selection. *Genetics* 168:1457–65.
- Swanson, W. J., Z. Yang, M. F. Wolfner, C. F. Aquadro. 2001b. Positive Darwinian selection drives the evolution of several female reproductive proteins in mammals. *P. Natl. Acad. Sci. USA* 98:2509–14.
- Tardif, S., N. Cormier. 2011. Role of zonadhesin during sperm-egg interaction: a species-specific acrosomal molecule with multiple functions. *Mol. Hum. Reprod.* 17:661–68.
- Tardif, S., M. D. Wilson, R. Wagner, P. Hunt, M. Gertsenstein, A. Nagy, C. Lobe, B. F. Koop, D. M. Hardy. 2010. Zonadhesin is essential for species specificity of sperm adhesion to the egg zona pellucida. *J. Biol. Chem.* 285:24863–70.
- Thompson, J. D., D. G. Higgins, T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acid. Res.* 22:4673–80.
- Thornhill, R., J. Alcock. 1983. *The Evolution of Insect Mating Systems*. Harvard Un. Cambridge.
- Torgerson, D. G., R. J. Kulathinal, R. S. Singh. 2002. Mammalian sperm proteins are rapidly evolving: evidence of positive selection in functionally diverse genes. *Mol. Biol. Evol.* 19:1973–80.
- Tourmente, M., M. Gomendio, E. R. S. Roldan. 2011a. Mass-specific metabolic rate and sperm competition determine sperm size in marsupial mammals. *PLoS One* 6:e21244.
- Tourmente, M., M. Gomendio, E. R. S. Roldan. 2011b. Sperm competition and the evolution of sperm design in mammals. *BMC Evol. Biol.* 11:12.
- Tourmente, M., E. R. S. Roldan, L. C. Giojalas, M. Chiaraviglio. 2009. Sperm competition and reproductive mode influence sperm dimensions and structure among snakes. *Evolution*. 63:2513–24.

- Tourmente, M., M. Rowe, M. M. González-Barroso, E. Rial, M. Gomendio, E. R. S. Roldan. 2013. Postcopulatory Sexual Selection Increases ATP Content in Rodent Spermatozoa. *Evolution*. DOI: 10.1111/evo.12079.
- Tsaur, S. C., C. I. Wu. 1997. Positive selection and the molecular evolution of a gene of male reproduction, Acp26Aa of *Drosophila*. *Mol. Biol. Evol.* 14:544–49.
- Turner, E., R. Montgomerie. 2002. Ovarian fluid enhances sperm movement in Arctic charr. *J. Fish Biol.* 60:1570–79.
- Turner, L. M., H. E. Hoekstra. 2006. Adaptive evolution of fertilization proteins within a genus: variation in ZP2 and ZP3 in deer mice (*Peromyscus*). *Mol. Biol. Evol.* 23:1656–69.
- Turner, L. M., H. E. Hoekstra. 2008a. Causes and consequences of the evolution of reproductive proteins. *Int. J. Dev. Biol.* 52:769–80.
- Turner, L. M., H. E. Hoekstra. 2008b. Reproductive protein evolution within and between species: maintenance of divergent ZP3 alleles in *Peromyscus*. *Mol. Ecol.* 17:2616–28.
- Turner, R. M. 2006. Moving to the beat : a review of mammalian sperm motility regulation. *Repr Fert Dev.* 18:25–38.
- Vacquier, V. D., G. W. Moy. 1977. Isolation of bindin: the protein responsible for adhesion of sperm to sea urchin eggs. *Proc. Natl. Acad. Sci. USA.* 74:2456–60.
- Vacquier, V. D., W. J. Swanson. 2011. Selection in the rapid evolution of gamete recognition proteins in marine invertebrates. *Cold Spring Harb. Perspect. Biol.* 3:a002931.
- Vanin, E. F. 1985. Processed Pseudogenes: Characteristics and Evolution. *Annu. Rev. Genet.* 19:253–72.
- Varea Sánchez, M., M. Bastir, E. R. S. Roldan. 2013. Geometric Morphometrics of Rodent Sperm Head Shape. *PLoS One* 8:e80607.
- Visconti, E., D. Moore. 1998. The Molecular Basis of Sperm Capacitation. *J. Androl.* 19:242–48.
- Voolstra, C., D. Tautz, P. Farbrother, L. Eichinger, B. Harr. 2007. Contrasting evolution of expression differences in the testis between species and subspecies of the house mouse. *Genome Res.* 17:42–9.
- Wagner, R. H., F. Helfenstein, E. Danchin. 2004. Female choice of young sperm in a genetically monogamous bird. *Proc. Biol. Sci.* 271:S134–37.
- Ward, P. I. 2000. Cryptic female choice in the yellow dung fly *Scathophaga stercoraria* (L.). *Evolution* 54:1680–86.
- Wasbrough, E. R., S. Dorus, S. Hester, J. Howard-Murkin, K. Lilley, E. Wilkin, A. Polpitiya, K. Petritis, T. L. Karr. 2010. The *Drosophila melanogaster* sperm proteome-II (DmSP-II). *J. Proteomics* 73:2171–85.
- Wassarman, P. M., L. Jovine, E. S. Litscher. 2001. A profile of fertilization in mammals. *Nat. Cell Biol.* 3:59–64.

- Wassarman, P. M., L. Jovine, E. S. Litscher. 2004. Mouse zona pellucida genes and glycoproteins. *Cytogenet. Genome Res.* 105:228–34.
- Welch, J. E., P. L. Brown, D. A. O. Brien, P. L. Magyar, D. O. Bunch, C. Mori, E. M. Eddy. 2000. Human Glyceraldehyde 3-Phosphate Dehydrogenase-2 Gene Is Expressed Specifically in Spermatogenic Cells. *Environ. Prot.* 21:328–38.
- Wilson, D. E., D. M. Reeder. 2005. *Mammal Species of the World: A Taxonomic and Geographic reference.* John Hopkins Univerisity Press. Baltimore.
- Wira, C. R., J. V Fahey, C. L. Sentman, P. A. Pioli, L. Shen. 2005. Innate and adaptive immunity in female genital tract: cellular responses and interactions. *Immunol. Rev.* 206:306–35.
- Wolf, J. B. W., A. Künstner, K. Nam, M. Jakobsson, H. Ellegren. 2009. Nonlinear dynamics of nonsynonymous (dN) and synonymous (dS) substitution rates affects inference of selection. *Genome Biol. Evol.* 2009:308–19.
- Wolfner, M. F. 1997. Mini-Review Tokens of Love : Functions and Regulation of Drosophila Male Accessory Gland Products. *Insect Biochem. Mol. Biol.* 27:179–92.
- Wolfner, M. F. 2002. The gifts that keep on giving: physiological functions and evolutionary dynamics of male seminal proteins in Drosophila. *Heredity* 88:85–93.
- Wong, A. 2010. Testing the effects of mating system variation on rates of molecular evolution in primates. *Evolution* 64:2779–85.
- Wong, A. 2011. The Molecular Evolution of Animal Reproductive Tract Proteins : What Have We Learned from Mating-System Comparisons ? *Int. J. Evol. Biol.* 2011:908735.
- Wyckoff, G. J., W. Wang, C. Wu. 2000. Rapid evolution of male reproductive genes in the descent of man. *Nature* 403:304–09.
- Yanagimachi, R. 1968. In vitro capacitation of hamster spermatozoa by follicular fluid. *J. Reprod. Fertil.* 18:275–86.
- Yang, Z. 1998. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. *Mol. Biol. Evol.* 15:568–73.
- Yang, Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24:1586–91.
- Yang, Z., J. P. Bielawski. 2000. Statistical methods for detecting molecular adaptation. *Trends Ecol. Evol.* 15:496–503.
- Yang, Z., R. Nielsen. 1998. Synonymous and nonsynonymous rate variation in nuclear genes of mammals. *J. Mol. Evol.* 46:409–18.
- Yang, Z., R. Nielsen, N. Goldman, A. Pedersen. 2000a. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics.* 155:431–49.

Yang, Z., W. J. Swanson, V. D. Vacquier. 2000b. Maximum-likelihood analysis of molecular adaptation in abalone sperm lysin reveals variable selective pressures among lineages and sites. *Mol. Biol. Evol.* 17:1446–55.

Zhang, J., R. Nielsen, Z. Yang. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol. Biol. Evol.* 22:2472–79.

Zhu, G., B. J. Miller, C. Boucheix, E. Rubinstein, C. C. Liu, R. O. Hynes, D. G. Myles, P. Primakoff. 2002. Residues SFQ (173-175) in the large extracellular loop of CD9 are required for gamete fusion. *Development* 129:1995–2002.

Zigler, K. S. 2008. The evolution of sea urchin sperm bindin. *Int. J. Dev. Biol.* 52:791–96.

Zigler, K. S., M. A. McCartney, D. R. Levitan, H. A. Lessios. 2005. Sea urchin bindin divergence predicts gamete compatibility. *Evolution* 59:2399–404.

Zigler, K. S., E. C. Raff, E. Popodi, R. A. Raff, H. A. Lessios. 2003. Adaptive evolution of bindin in the genus *Heliocidaris* is correlated with the shift to direct development. *Evolution* 57:2293–302.

Ziyyat, A., E. Rubinstein, F. Monier-gavelle, V. Barraud, O. Kulski, M. Prenant, C. Boucheix, M. Bomsel, J. Wolf. 2006. CD9 controls the formation of clusters that contain tetraspanins and the integrin alpha 6 beta 1, which are involved in human and mouse gamete fusion. *J. Cell Sci.* 119:416–24.